PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 5 :
C12N 15/31, 9/80, C12Q 1/68, C12P 21/08, A61K 39/106, G01N 33/577

A1

(11) International Publication Number: **WO 94/26901**

(43) International Publication Date: 24 November 1994 (24.11.94)

(21) International Application Number: PCT/EP94/01625

(22) International Filing Date: 19 May 1994 (19.05.94)

(30) Priority Data:
93401309.5 19 May 1993 (19.05.93) EP
(34) Countries for which the regional or international application was filed: GB et al.
PCT/EP93/03259 19 November 1993 (19.11.93) WO
(34) Countries for which the regional or international application was filed: JP et al.

(71) Applicants *(for all designated States except US)*: INSTITUT PASTEUR [FR/FR]; 25-28, rue du Dr.-Roux, F-75724 Paris Cédex 15 (FR). INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE [FR/FR]; 101, rue de Tolbiac, F-75654 Paris Cédex 13 (FR).

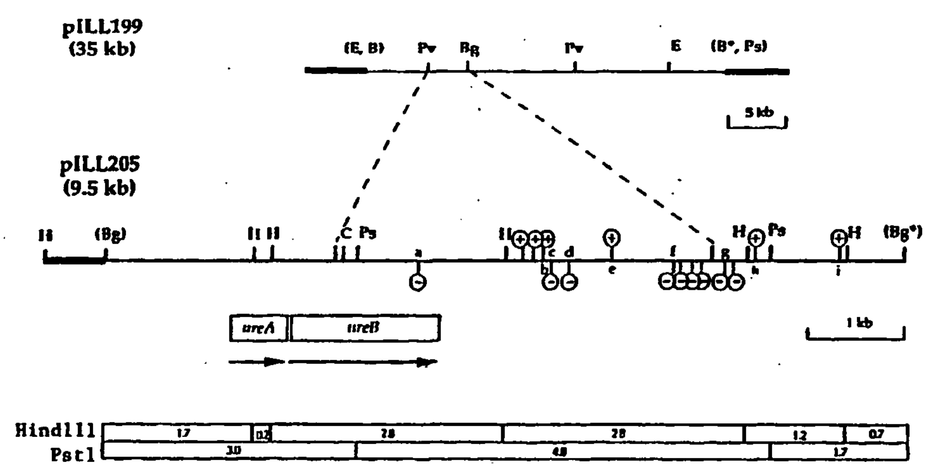
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LABIGNE, Agnès [FR/FR]; 47, avenue Beauséjour, F-91440 Bures-sur-Yvette (FR). SUERBAUM, Sébastien [FR/FR]; 40, rue Spontini, F-75116 Paris (FR). FERRERO, Richard [FR/FR]; 60, avenue des Gobelins, F-75013 Paris (FR). THIBERGE, Jean-Michel [FR/FR]; Appartement 532, 15, rue de la Ferronnerie, F-78270 Plaisir (FR).

(74) Agents: GUTMANN, Ernest et al.; Ernest Gutmann-Yves Plasseraud S.A., 3, rue Chauveau-Lagarde, F-75008 Paris (FR).

(81) Designated States: AU, CA, JP, KR, US, European patent (AT, BE, CH, DE, DK, ES, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: IMMUNOGENIC COMPOSITIONS AGAINST HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE COMPOSITIONS AND NUCLEIC ACID SEQUENCES ENCODING SAID POLYPEPTIDES



(57) Abstract

The invention relates to an immunogenic composition, capable of inducing protective antibodies against *Helicobacter* infection, characterised in that it comprises: i) at least one sub-unit of a urease structural polypeptide from *Helicobacter pylori*, or a fragment thereof, said fragment being recognised by antibodies reacting with *Helicobacter felis* urease, and/or at least one sub-unit of a urease structural polypeptide from *Helicobacter felis*, or a fragment thereof, said fragment being recognised by antibodies reacting with *Helicobacter pylori* urease; ii) and/or, a Heat Shock protein (HSP), or chaperonin, from *Helicobacter*, or a fragment of said protein. The invention also relates to the preparation, by recombinant means, of such immunogenic compositions.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AT	Austria
AU	Australia
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CS	Czechoslovakia
CZ	Czech Republic
DE	Germany
DK	Denmark
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgystan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakhstan
LI	Liechtenstein
LK	Sri Lanka
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SI	Slovenia
SK	Slovakia
SN	Senegal
TD	Chad
TG	Togo
TJ	Tajikistan
TT	Trinidad and Tobago
UA	Ukraine
US	United States of America
UZ	Uzbekistan
VN	Viet Nam

IMMUNOGENIC COMPOSITIONS AGAINST HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE COMPOSITIONS AND NUCLEIC ACID SEQUENCES ENCODING SAID POLYPEPTIDES

The present invention relates to immunogenic compositions for inducing protective antibodies against Helicobacter spp. infection. It also relates to proteinaceous material derived from Helicobacter, and to nucleic acid sequences encoding them. Antibodies to these proteinaceous materials are also included in the invention.

H. pylori is a microorganism which infects human gastric mucosa and is associated with active chronic gastritis. It has been shown to be an aetiological agent in gastroduodenal ulceration (Peterson, 1991) and two recent studies have reported that persons infected with H. pylori had a higher risk of developing gastric cancer (Nomura et al, 1991 ; Parsonnet et al, 1991).

In vivo studies of the bacterium and, consequently, work on the development of appropriate preventive or therapeutic agents has been severely hindered by the fact that Helicobacter pylori only associates with gastric-type epithelium from very few animal hosts, none of which are suitable for use as laboratory models.

A mouse model of gastric colonisation has been developed using a helical bacterium isolated from cat gastric mucus (Lee et al, 1988, 1990) and identified as a member of the genus Helicobacter. It has been named H. felis (Paster et al, 1990).

To date, only limited information concerning H. felis and the extent of its similarities and

differences with H. pylori, is available. The reliability of the mouse model for the development of treatments for H. pylori infection is therefore uncertain. Recently, it was shown that H. pylori urease is a protective antigen in the H. felis / mouse model (Davin et al, 1993 ; Corthesy-Theulaz et al, 1993).

It is therefore an aim of the present invention to provide therapeutic and preventive compositions for use in Helicobacter infection, which furthermore can be tested in laboratory animals.

It is known that H. pylori expresses urease activity and that urease plays an important role in bacterial colonisation and mediation of certain pathogenic processes (Ferrero and Lee, 1991 ; Hazel et al, 1991).

The genes coding for the urease structural polypeptides of H. pylori (URE A, URE B) have been cloned and sequenced (Labigne et al, 1991 ; and French Patent Application FR 8813135), as have the genes coding the "accessory" polypeptides necessary for urease activity in H. pylori (International patent application WO 93/07273).

Attempts have been made to use nucleic acid sequences from the H. pylori urease gene cluster as probes to identify urease sequences in H. felis. However, none of these attempts have been successful. Furthermore, the establishment and maintenance of H. felis cultures in vitro is extremely difficult, and the large quantities of nucleases present in the bacteria complicates the extraction of DNA.

The present inventors have however, succeeded in cloning and sequencing the genes of the urease structural polypeptides of H. felis, and of the accessory polypeptides. This has enabled, in the

context of the invention, the comparison of the amino-acid sequence data for the H. felis ure gene products with that for Helicobacter pylori, and a high degree of conservation between the urease sub-units has been found. An immunological relationship between the 2 ureases exists, and protective antibodies to Helicobacter infection can be induced using the urease sub-units or fragments thereof as immunogens.

Indeed, to elucidate the efficiency of individual urease subunits to act as mucosal immunogens, the genes encoding the respective urease subunits (UreA and UreB) of Helicobacter pylori and Helicobacter felis have been cloned in an expression vector (pMAL), and expressed in Escherichia coli cells as translational fusion proteins. The recombinant UreA and UreB proteins have been purified by affinity and anion exchange chromatography techniques, and have predicted molecular weights of approximately 68 and 103 kDa, respectively. Western blotting studies indicated that the urease components of the fusion proteins are strongly immunogenic and are specifically recognized by polyclonal rabbit anti-Helicobacter sera. Orogastric immunization of mice with 50 µg of recombinant H. felis UreB, administered in combination with a mucosal adjuvant (cholera toxin), protected 60 % ($n = 7$; $p < 0.005$) of mice from gastric colonization by H. felis bacteria at over 4 months. This compared with a value of 25 % ($n = 8$; $p > 0.05$) for the heterologous H. pylori UreB antigen. For the first time, a recombinant subunit antigen has been shown to induce an immunoprotective response against gastric Helicobacter infection.

The inventors have also identified, in the context of the invention, new Heat Shock Proteins or chaperonins, in Helicobacter, which have an enhancing

effect on urease activity. Use of the chaperonins in an immunogenic composition may induce therefore an enhancement of protection.

Indeed, the genes encoding each of the HspA and HspB polypeptides of Helicobacter pylori have been cloned, expressed independently as fused proteins to the Maltose-Binding-Protein (MBP), and purified on a large scale. These proteins have been used as recombinant antigens to immunize rabbits, and in Western immunoblotting assays as well as ELISA to determine their immunogenicity in patients infected with HP (HP+). The MBP-HspA and MBP-HspB fusion proteins have been shown to retain their antigenic properties. Comparison of the humoral immune response against HspA and/or HspB in (HP+) patient sera demonstrated that not only HspB but also HspA was recognized by (HP+) patient sera (29/38 and 15/38, respectively). None of the 14 uninfected patients had antibodies reacting with the Hsps.

The present invention concerns an immunogenic composition capable of inducing antibodies against Helicobacter infection characterised in that it comprises :

i) at least one sub-unit of a urease structural polypeptide from Helicobacter pylori, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter felis urease, and/or at least one sub-unit of a urease structural polypeptide from Helicobacter felis, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter pylori urease ;

ii) and/or a Heat Shock protein (HSP), or chaperonin, from Helicobacter, or a fragment of said protein.

Preferably, the immunogenic composition is capable of inducing protective antibodies.

According to a preferred embodiment, the immunogenic composition of the invention contains, as the major active ingredient, at least one sub-unit of a urease structural polypeptide from Helicobacter pylori and/or Helicobacter felis. The expression "urease structural polypeptide" signifies, in the context of the present invention, the enzyme of Helicobacter pylori or Helicobacter felis probably a major surface antigen composed of two repeating monomeric sub-units, a major sub-unit (product of the ure B gene) and a minor sub-unit, product of the ure A gene and which, when complemented by the presence of the products of the accessory genes of the urease gene cluster, are responsible for urease activity i.e. the hydrolysis of urea to liberate NH_4^+ in the two Helicobacter species. It is to be understood that in the absence of the accessory gene products, the urease structural polypeptides do not exhibit enzymatic activity, but are recognised by antibodies reacting with H. felis or H. pylori urease.

The term "immunogenic composition" signifies, in the context of the invention, a composition comprising a major active ingredient as defined above, together with any necessary ingredients to ensure or to optimise an immunogenic response, for example adjuvants, such as mucosal adjuvant, etc...

The Helicobacter pylori urease structural polypeptide has been described and sequenced by Labigne et al, 1991. The polypeptide described in this paper is particularly appropriate for use in the composition of the present invention. However, variants showing functional homology with this published sequence may be used, which comprise amino-

acid substitutions, deletions or insertions provided that the immunological characteristics of the polypeptide in so far as its cross-reactivity with anti-Helicobacter felis urease antibodies is concerned, are maintained. Generally speaking, the polypeptide variant will show a homology of at least 75% and preferably about 90% with the included sequence.

A fragment of the Helicobacter pylori urease structural polypeptide may also be used in the immunogenic composition of the invention, provided that the fragments are recognised by antibodies reacting with Helicobacter felis urease. Such a fragment will generally be comprised of at least 6 amino-acids, for example, from 6 to 100 amino-acids, preferably about 20-25. Advantageously, the fragment carries epitopes unique to Helicobacter.

Nucleic acid and amino-acid sequences may be interpreted in the context of the present invention by reference to figures 11 and 12, showing the genetic code and amino-acid abbreviations respectively.

The Helicobacter felis urease structural polypeptide suitable for use in the present invention is preferably that encoded by part of the plasmid pILL205 (deposited at the CNCM on 25th August 1993, under number : CNCM I-1355), and whose amino-acid sequence is shown in figure 3 (subunits A and B). Again, a variant of this polypeptide comprising amino-acid substitutions, deletions or insertions with respect to the figure 3 sequence may be used provided that the immunological cross-relationship with Helicobacter pylori urease is maintained. Such a variant normally exhibits at least 90 % homology or identity with the figure 3 sequence. An example of such variants are the urease A and B sub-units from

Helicobacter heilmannii (Solnick et al, 1994), shown to have 80 % and 92 % identity with the H. felis urease A and B sub-units, respectively.

Fragments of this urease or variants may be used in the immunogenic composition provided that the fragments are recognised by antibodies reacting with Helicobacter pylori urease. Again, the length of such a fragment is usually at least 6 amino-acids, for example from 6 to 100, preferably about 20 to 25. Preferably, the fragment carries epitopes unique to Helicobacter.

If variants or fragments of the native urease sequences are employed in the immunogenic composition of the invention, their cross-reactivity with antibodies reacting with urease from the other Helicobacter species can be tested by contacting the fragment or the variant with antibodies, preferably polyclonal raised to either the native or the recombinant urease or, alternatively, to whole Helicobacter. Preferably, the variants and fragments give rise to antibodies which are also capable of reacting with H. heilmannii urease. Cross protection to infection by H. heilmannii is therefore also obtained by the immunogenic composition of the invention.

The use of fragments of the urease structural genes is particularly preferred since the immunological properties of the whole polypeptide may be conserved whilst minimizing risk of toxicity.

The active component of the immunogenic composition of the invention may be comprised of one sub-unit only of the urease structural polypeptide, that is either sub-unit A or sub-unit B products of the ure A and ure B genes respectively. Compositions comprising only the urease sub-unit Ure B, of either

H. pylori or H. felis, or variants and fragments as defined above, are particularly advantageous. Most preferred are homologous systems wherein the urease sub-unit particularly sub-unit B, is derived from the organism against which protection is sought, e.g. H. felis sub-unit B against H. felis infection. However, the composition may contain both A and B sub-units, which are normally present as distinct polypeptides. However, it is possible, when the polypeptide is produced by recombinant means, to use a fusion protein comprising the entire sequences of the A and B gene products by the suppression of the stop-codon separating the two adjacent coding sequences.

The urease component of the immunogenic composition, whether sub-unit A or sub-unit B, may be used in the form of translational fusion proteins, for example with the Maltose-Binding-Protein (MBP). Other suitable fusions are exemplified in International Patent Application WO 90/11360. Another example of a suitable fusion protein is the "QIAexpress" system commercialised by QIAGEN, USA, which allows the 6xHis tag sequence to be placed at the 5' or 3' end of the protein coding sequence. The use of the active ingredients in the form of fusion proteins is however, entirely optional.

According to a further preferred embodiment, the immunogenic composition of the invention may comprise in addition to or instead of the urease structural polypeptide defined above, a Heat Shock Protein also known as a "chaperonin" from Helicobacter. These chaperonins have been elucidated by the inventors in the context of the present invention. Preferably, the chaperonin is from Helicobacter pylori. Such an HSP may be the urease-associated HSP A or HSP B or a mixture of the two, having the amino-acid sequence

illustrated in figure 6. These polypeptides are encoded by the plasmid pILL689 (deposited at CNCM on 25th August 1993, under number : CNCM I-1356). Particularly preferred is the H. pylori HSP-A protein, either alone or in combination with Hsp-B.

It is also possible to use, as HSP component, according to the invention, a polypeptide variant in which amino-acids of the figure 6 sequence have been replaced, inserted or deleted, the said variant normally exhibiting at least 75 %, and preferably at least 85 % homology with the native HSP. The variants preferably exhibit at least 75 %, for example at least 85 % identity with the native Hsp.

The variants may further exhibit functional homology with the native polypeptide. In the case of the HSP components, "functional homology" means the capacity to enhance urease activity in a micro-organism capable of expressing active urease, and/or the capacity to block infection by Helicobacter, particularly H. felis and H. pylori. The property of enhancing urease activity may be tested using the quantitative urease activity assay described below in the examples. Fragments of either or both of the HSP A and HSP B polypeptides preferably having at least 6 amino-acids, may be used in the composition. The fragments or variants of the HSP component used in the immunogenic composition of the invention are preferably capable of generating antibodies which block the urease enhancing effect normally exhibited by the HSPs. This property is also tested using the quantitative assay described in the examples. The presence of the chaperonins in the composition enhances the protection against Helicobacter pylori and felis.

The Hsp component of the immunogenic composition, whether HspA or HspB can be used in the form of a translational fusion protein, for example with the Maltose-Binding-Protein (MBP). As for the urease component, other suitable fusion partners are described in International Patent Application WO 90/11360. The "QIAexpress" system of QIAGEN, USA, may also be used. Again, the use of the proteins in the form of fusion proteins is entirely optional.

According to the invention therefore the immunogenic composition may comprise either a urease structural polypeptide as defined above, or a Helicobacter Hsp, particularly HspA or a combination of these immunogens.

According to a preferred embodiment, the immunogenic composition comprises, as urease component, both the A and B sub-units of both Helicobacter felis (i.e. without H. pylori urease) together with the HSP A and HSP B of Helicobacter pylori. Alternatively, the A and B sub-units of the Helicobacter felis urease may be used together with those of H. pylori, but without chaperonin component.

The immunological cross-reactivity between the ureases of the two different Helicobacter species enables the use of one urease only in the composition, preferably that of Helicobacter felis. The protective antibodies induced by the common epitopes will however be active against both Helicobacter pylori and Helicobacter felis. It is also possible that the composition induce protective antibodies to other species of Helicobacter, if the urease polypeptide or fragment carries epitopes occuring also on those other species.

The composition of the invention is advantag ously us d as an immunogenic compositi n or a

vaccine, together with physiologicaly acceptable excipients and carriers and, optionally, with adjuvants, haptens, carriers, stabilizers, etc. Suitable adjuvants include muranmyl dipeptide (MDP), complete and incomplete Freund's adjuvants (CFA and IFA) and alum. The vaccine compositions are normally formulated for oral administration.

The vaccines are preferably for use in man, but may also be administered in non-human animals, for example for vetinary purposes, or for use in laboratory animals such as mice, cats and dogs.

The immunogenic compositions injected into animals raises the synthesis <u>in vivo</u> of specific antibodies, which can be used for therapeutic purposes, for example in passive immunity.

The invention also relates to the proteinaceous materials used in the immunogenic composition and to proteinaceous material encoded by the urease gene clusters other than the A and B urease structural sub-units. "Proteinaceous material" means any molecule comprised of chains of amino-acids, eg. peptides, polypeptides or proteins, fusion or mixed proteins (i.e. an association of 2 or more proteinaceous materials, all or some of which may have immunogenic or immunomodulation properties), either purified or in a mixture with other proteinaceous or non-proteinaceous material. "Polypeptide" signifies a chain of amino-acids whatever its length and englobes the term "peptide". The term "fragment" means any amino-acid sequence shorter by at least one amino-acid than the parent sequence and comprising a length of amino-acids e.g. at least 6 residues, consecutive in the parent sequence.

The peptide sequences of the invention, may for example, be obtained by chemical synthesis, using a

technique such as the Merrifield technique and synthesiser of the type commercialised by Applied Biosystems.

In particular, the invention relates to proteinaceous material characterised in that it comprises at least one of the Helicobacter felis polypeptides encoded by the urease gene cluster of the plasmid pILL205 (CNCM I-1355), including the structural and accessory urease polypeptides, or a polypeptide having at least 90 % homology with said polypeptides, or a fragment thereof. Of particular interest are the gene products of the ure A and ure B genes, as illustrated in figure 3, or a variant thereof having at least 90 % homology or a fragment having at least 6 amino-acids. The fragments and the variants are recognised by antibodies reacting with Helicobacter pylori urease.

Amongst the polypeptides encoded by the accessory genes of the urease gene cluster, is the gene product of ure I, as illustrated in figure 9, which also forms part of the invention. Also included is a variant of the ure I product having at least 75 % homology, preferably at least 85 %, or a fragment of the gene product or of the variant having at least 6 amino-acids. The variant preferably has the capacity to activate the ure A and ure B gene products in the presence of the remaining urease accessory gene products. This functional homology can be detected by using the following test : 10^9 bacteria containing the ure I gene product variant are suspended in 1 ml of urea-indole medium and incubated at 37° C. The hydrolysis of the urea leads to the release of ammonium, which increases pH and induces a colour change from orange to fuscia-red. The observation of such a colour change demonstrates that the variant of

the ure I gene product under test is capable of activating the ure A and B gene products.

It is also possible that a fragment of the ure I gene product, if it has a length of, for example, at least 70 or 100 amino-acids, may also exhibit this functional homology with the entire polypeptide.

The fragments of ure I polypeptide or of the variant preferably are capable of inducing the formation of antibodies which block the urease maturation process. In other words, the fragments bear epitopes which play a decisive role in the interaction between the ure I and ure A / ure B gene products.

The invention also relates to the proteinaceous material comprising at least one of the Heat Shock Proteins or chaperonins of Helicobacter pylori or a fragment thereof. Particularly preferred are the HSP A and HSP B polypeptides as illustrated in figure 6 or a polypeptide having at least 75 %, and preferably at least 80 or 90 %, homology or identity with the said polypeptide. A particularly preferred fragment of the Helicobacter pylori HSP A polypeptide is the C-terminal sequence :

G S C C H T G N H D H K H A K E H E A C C H D H K K H

or a sub-fragment of this sequence having at least 6 consecutive amino-acids. This C-terminal sequence is thought to act as a metal binding domain allowing binding of, for example, nickel.

The proteinaceous material of the invention may also comprise or consist of a fusion or mixed protein including at least one of the sub-units of the urease structural polypeptide of H. pylori and/or of H. felis, or fragments or variants thereof as defined above. Particularly preferred fusion proteins are the

Mal-E fusion proteins and QIAexpress system fusion proteins (QIAGEN, USA) as detailed above. The fusion or mixed protein may include, either instead of in addition to the urease sub-unit, a Heat Shock Protein, or fragment or variant thereof, as defined above.

The invention also relates to monoclonal or polyclonal antibodies to the proteinaceous materials described above. More particularly, the invention relates to antibodies or fragments thereof to any one of the Helicobacter felis polypeptides encoded by the urease gene cluster of the plasmid pILL205 (CNCM I-1355) including the structural and accessory urease polypeptides that is, structural genes ure A and ure B and the accessory genes known as ure C, ure D, ure E, ure F, ure G, ure H and ure I. The antibodies may also be directed to a polypeptide having at least 90 % homology with any of the above urease polypeptides or to a fragment thereof preferably having at least 6 amino-acids. The antibodies of the invention may specifically recognise Helicobacter felis polypeptides expressed by the urease gene cluster. In this case, the epitopes recognised by the antibodies are unique to Helicobacter felis. Alternatively, the antibodies may include or consist of antibodies directed to epitopes common to Helicobacter felis urease polypeptides and to Helicobacter pylori urease polypeptides. If the antibodies recognise the accessory gene products, it is particularly advantageous that they cross-react with the Helicobacter pylori accessory gene product. In this way, the antibodies may be used in therapeutic treatment of Helicobacter pylori infection in man, by blocking the urease maturation process.

Particularly preferred antibodies of the invention recognise the Helicobacter felis ure A

and/or ure B gene products, that is the A and B urease sub-units. Advantageously, these antibodies also cross-react with the Helicobacter pylori A and B urease sub-units, but do not cross-react with other ureolytic bacteria. Such antibodies may be prepared against epitopes unique to Helicobacter (see figure 4), or alternatively, against the whole polypeptides followed by screening out of any antibodies reacting with other ureolytic bacteria.

The invention also concerns monoclonal or polyclonal antibodies to the HSPs or fragments thereof, particularly to the HSP A and/or HSP B protein illustrated in figure 6. Polypeptides having at least 75 %, and preferably at least 80 %, or 90 % homology with the HSPs may also be used to induce antibody formation. These antibodies may be specific for the Helicobacter pylori chaperonins or, alternatively, they may cross-react with GroEL-like proteins or GroES-like proteins from bacteria other than Helicobacter, depending upon the epitopes recognised. Figure 7 shows the homologous regions of HSP A and HSP B with GroES-like proteins and GroEL-like proteins respectively from various bacteria. Particularly preferred antibodies are those specific for either the HSP A or HSP B chaperonins or those specifically recognising the HSP A C-terminal sequence having the metal binding function. Again, use of specific fragments for the induction of the antibodies ensures production of Helicobacter-specific antibodies.

The antibodies of the invention may be prepared using classical techniques. For example monoclonal antibodies may be produced by the hybridoma technique or by known techniques for the preparation of human antibodies, or by the technique described by Marks et

al (Journal of Molecular Biology, 1991, 222, p 581-597).

The invention also includes fragments of any of the above antibodies produced by enzyme digestion. Of particular interest are the Fab and $F(ab')_2$ fragments. Also of interest are the Facb fragments.

The invention also relates to purified antibodies or serum obtained by immunisation of an animal, e.g. a mammal, with the immunogenic composition, the proteinaceous material or fragment, or the fusion or mixed protein of the invention, followed by purification of the antibodies or serum. Also concerned is a reagent for the in vitro detection of H. pylori infection, containing at least these antibodies or serum, optionally with reagents for labelling the antibodies e.g. anti-antibodies etc.

The invention further relates to nucleic acid sequences coding for any of the above proteinaceous materials including peptides. In particular, the invention relates to a nucleic acid sequence characterised in that it comprises :

i) a sequence coding for the Helicobacter felis urease and accessory polypeptides as defined above, and a sequence coding for the HSP of H. pylori as defined above ;

or ii) a sequence complementary to sequence (i) ;

or iii) a sequence capable of hybridizing to sequence (i) or (ii) under stringent conditions ;

or iv) a fragment of any of sequences (i), (ii) or (iii) comprising at least 10 nucleotides.

Preferred nucleic acid sequences are those comprising all or part of the sequence of plasmid pIL205 (CNCM I-1355), for example the sequence of Figure 3, in particular that coding for the gene product of ure A and for ure B or the sequence of

Figure 9 (<u>Ure I</u>), or a sequence capable of hybridising with these sequences under stringent conditions, or a sequence complementary to these sequences, or a fragment comprising at least 10 consecutive nucleotides of these sequences.

Other preferred sequences are those comprising all or part of the sequence of plasmid pILL689 (CNCM I-1356), for example the sequence of figure 6, in particular that coding for HSP A and/or HSP B, or a sequence complementary to this sequence, or a sequence capable of hybridizing to this sequence under stringent conditions, or a fragment thereof.

High stringency hybridization conditions in the context of the invention are the following :

- 5 x SSC ;
- 50 % formamide at 37°C ;

or :

- 6 x SSC ;
- Denhard medium at 68°C.

The sequences of the invention also include those hybridizing to any of sequences (i), (ii) and (iii) defined above under non-stringent conditions, that is :

- 5 x SSC ;
- 0.1 % SDS ;
- 30 or 40 % formamide at 42°C, preferably 30 %.

The term "complementary sequences" in the context of the invention signifies "complementary" and "reverse" or "inverse" sequences.

The nucleic acid sequences may be DNA or RNA.

The sequences of the invention may be used as nucleotide probes in association with appropriate labelling means. Such means include radio-active isotopes, enzymes, chemical or chemico-luminescent markers, fluoro-chromes, haptens, or antibodies. The

markers may optionally be fixed to a solid support, for example a membrane, or particles.

As a preferred marker, radio-active phosporous (^{32}P) is incorporated at the 5'-end of the probe sequence. The probes of the invention comprise any fragment of the described nucleic acid sequences and may have a length for example of at least 45 nucleotides, for example 60, 80 or 100 nucleotides or more. Preferred probes are those derived from the ure A, ure B, ure I, HSP A and HSP B genes.

The probes of the invention may be used in the in vitro detection of Helicobacter infection in a biological sample, optionally after a gene amplification reaction. Most advantageously, the probes are used to detect Helicobacter felis or Helicobacter pylori, or both, depending on whether the sequence chosen as the probe is specific to one or the other, or whether it can hybridise to both. Generally, the hybridisation conditions are stringent in carrying out such a detection.

The invention also relates to a kit for the in vitro detection of Helicobacter infection, characterised in that it comprises :

- a nucleotide probe according to the invention, as defined above ;
- an appropriate medium for carrying out a hybridisation reaction between the nucleic acid of Helicobacter and the probe ;
- reagents for the detection of any hybrids formed.

The nucleotide sequences of the invention may also serve as primers in a nucleic acid amplification reaction. The primers normally comprise at least 10 consecutive nucleotides of the sequences described above and preferably at least 18. Typical lengths are

from 25 to 30 and may be as high as 100 or more consecutive nucleotides. Such primers are used in pairs and are chosen to hybridize with the 5' and 3'-ends of the fragment to be amplified. Such an amplification reaction may be performed using for example the PCR technique (European patent applications EP200363, 201184 and 229701). The Q-β-replicase technique (Biotechnology, vol. 6, Oct. 1988) may also be used in the amplification reaction.

The invention also relates to expression vectors characterised in that they contain any of the nucleic acid sequences of the invention. Particularly preferred expression vectors are plasmids pILL689 and pILL205 (CNCM I-1356 and CNCM I-1355, respectively). The expression vectors will normally contain suitable promoters, terminators and marker genes, and any other regulatory signals necessary for efficient expression.

The invention further relates to prokaryotic or eukaryotic host cells stably transformed by the nucleic acid sequences of the invention. As examples of hosts, mention may be made of higher eukaryotes such as CHO cells and cell-lines ; yeast, prokaryotes including bacteria such as E. coli e.g E. coli HB 101 ; Mycobacterium tuberculosum ; viruses including baculovirus and vaccinia. Usually the host cells will be transformed by vectors. However, it is also possible within the context of the invention, to insert the nucleic acid sequences by homologous recombination, using conventional techniques.

By culturing the stably transformed hosts of the invention, the Helicobacter urease polypeptide material and, where applicable, the HSP material can be produced by recombinant means. The recombinant proteinaceous materials are then collected and purified. Pharmaceutical compositions are prepared by

combining th recombinant materials with suitable excipients, adjuvants and optionally, any other additives such as stabilizers.

The invention also relates to plasmids pILL920 (deposited at CNCM on 20.07.1993, under accession number I-1337) and pILL927 (CNCM I-1340, deposited on 20.07.1993) constructed as described in the examples below.

Different aspects of the invention are illustrated in the figures :

Figure 1 :

Transposon mutagenesis and sequencing of pILL205. Linear restriction maps of recombinant cosmid pILL199 and recombinant plasmid pILL205 (and the respective scale markers) are presented. Numbers in parentheses indicate the sizes of H.felis DNA fragments inserted into one of the cloning vectors (pILL575 or pILL570, respectively). The "plus" and "minus" signs within circles correspond to the insertion sites of the MiniTn3-Km transposon in pILL205 ; "plus" signs indicate that the transposon did not inactivate urease expression, whereas negative signs indicate that urease expression was abolished. The letters refer to mutant clones which were further characterised for quantitative urease activity and for the synthesis of urease gene products. The location of the structural urease genes (ure A and ure B) on pILL205 are represented by boxes, the lengths of which are proportional to the sizes of the respective open-reading frames. The arrows refer to the orientation of transcription. The scale at the bottom of the figure indicates the sizes (in kilobases) of the HindIII and PstI restriction fragments. Restriction sites are

represented as follows : B, BamHI ; Pv, PvuII ; Bg, BglII ; E, EcoRI ; H, HindIII ; C, ClaI ; Ps, PstI. Letters within parentheses indicate that the sites originated from the cloning vector.

Figure 2 :

Western blot analysis of whole-cell extracts of E. coli HB101 cells harbouring recombinant plasmids were reacted with rabbit polyclonal antiserum (diluted 1:1, 1000) raised against H. felis bacteria. A) extracts were of E. coli cells harbouring : plasmid vector pILL570 (lane 1) ; recombinant plasmid pILL205 (lane 2) ; and pILL205 derivative plasmids disrupted in loci "a", "b", "c", "d", and "e" (lanes 3-7). B) Extracts were of E. coli cells harbouring : recombinant plasmid pILL753 containing the H. pylori ure A and ure B genes (Labigne et al., 1991) (lane 1) ; and pILL205 derivative plasmids disrupted in loci "f", "g", "h", and "i" (lanes 2-5). The small arrow heads indicate polypeptides of approximately 30 and 66 kilodaltons which represent putative Ure A and Ure B gene products of H. felis. The large arrow heads in panel B indicate the corresponding gene products of H. pylori which cros-reacted with the anti-H. felis serum. The numbers indicate the molecular weights (in thousands) of the protein standards.

Figure 3 :

Nucleotide sequence of the H. felis structural urease genes. Numbers above the sequence indicate the nucleotide positions as well as the amino acid position in each of the two Ure A and Ure B polypeptides. Predicted amino acid sequences for Ure A (bp 43 to 753) and Ure B (766 to 2616) are shown below

the sequence. The putative ribosome-binding site (Shine-Dalgarno sequence, SD) is underlined.

Figure 4 :

Comparison of sequences for the structural urease genes of H. felis to : a) the sequence of the two subunits of H. pylori urease (Labigne et al., 1991) ; b) the sequence of the three subunits of Proteus mirabilis urease (Jones and Mobley, 1989) ; c) the sequence of the single subunit of jack bean urease. Gaps (shown by dashes) have been introduced to ensure the best alignment. *, amino acids identical to those of the H. felis sequence ; =, amino-acids shared by the various ureases ; ·, amino-acids unique to the Helicobacter ureases. The percentages relate to the number of amino acids that are identical to those of the H. felis urease subunits. H.f., Helicobacter felis ; H.p., Helicobacter pylori ; P.m., Proteus mirabilis ; J.b., Jack bean.

Figure 5 :

Restriction map of the recombinant plasmids pILL689, pILL685, and pILL691. The construction of these plasmids is described in details in Table 1. Km within triangles depictes the site of insertion of the kanamycin cassette which led to the construction of plasmids pILL687, pILL688 and pILL696 (table 2). Boxes underneath the maps indicate the position of the three genetic elements deduced from the nucleotide sequence, namely IS5, Hsp A and Hsp B.

Figure 6 :

Nucleotide sequence of the Helicobacter pylori Heat Shock Protein gene cluster. The first number above the sequence indicates the nucleotide positi ns, whereas the second one numbers the amino-acid residue

position for each of the Hsp A and Hsp B protein. The putative ribosome-binding sequences (Shine- Dalgarno [SD] sites) are underlined.

Figure 7 :

Comparison of the deduced amino-acid sequence of Helicobacter pylori Hsp A (A) or Hsp B (B) with that of other GroEL-like (A) or GroES-like (B) proteins. Asteriks mark amino-acids identical with those in the Helicobacter pylori Hsp A or Hsp B sequences.

Figure 8 :

Expression of the Helicobacter pylori Hsp A Heat-Shock proteins in E. coli minicells. The protein bands with apparent molecular masses of 58 and 13 kDA, corresponding to the Helicobacter pylori Hsp A and Hsp B Heat-Shock Proteins are clearly visible in the lanes corresponding to plasmids pILL689 and pILL692 and absent in the vector controls (pILL570 and pACYC177, respectively)

Figure 9 :

Nucleotide sequence of the Helicobacter felis ure I gene and deduced amino-acid sequence.

Figure 10 :

Comparison of the amino-acid sequence of the ure I proteins deduced from the nucleotide sequence of the ure I gene of Helicobacter felis and that of Helicobacter pylori.

Figure 11 :

Genetic code. Chain-terminating, or "nonsense", codons. Also used to specify the initiator formyl-Met-$tRNA^{Met}_{F}$. The Val triplet GUG is therefore

"ambiguous" in that it codes both valine and methionine.

Figure 12 :

Signification of the one-letter and three-letter amino-acid abbreviations.

Figure 13 :

Purification of H. pylori UreA-MBP recombinant protein using the pMAL expression vector system. Extracts from the various stages of protein purification were migrated on a 10 % resolvving SDS-polyacrylamide gel. Following electrophoresis, the gel was stained with Coomassie blue. The extracts were : 1) non-induced cells ; 2) IPTG-induced cells ; French press lysate of induced cell extract ; 5) eluate from amylose resin column ; 6) eluate from anion exchange column (first passage) ; 7) eluate from anion exchange column (second passage) ; 8) SDS-PAGE standard marker proteins.

Figure 14 :

Recognition of UreA recombinant fusion proteins by polyclonal rabbit anti-Helicobacter sera. Protein extracts of maltose-binding protein (MBP, lane 1), H. felis UreA-MBP (lane 2), and H. pylori UreA-MBP (lane 3) were Western Blotted using rabbit polyclonal antisera (diluted 1 : 5000) raised against whole-cell extracts of H. pylori and H. felis. The purified fusion proteins are indicated by an arrow. Putative degradation products of the proteins are shown by an asterisk.

Figure 15 :

Recognition of UreB recombinant fusion proteins by rabbit antisera raised against purified homologous

and heterologous UreB proteins. Nitrocellulose membranes were blotted with the following extracts : 1) standard protein markers ; 2) H. felis UreA-MBP ; 3) MBP ; 4) H. pylori UreA-MBP. The membranes were reacted with polyclonal rabbit antisera (diluted 1 : 5000) raised against MBP-fused H. pylori and H. felis Ure B sub-units, respectively. The molecular weights of standard proteins are presented on the left-hand side of the blots.

Figure 16 :

Western blot analysis of H. pylori and H. felis whole-cell extracts with antisera raised against purified UreB MBP-fused recombinant proteins. SDS-PAGE whole extracts of H. Felis (lane 1) and H. pylori (lane 2) cells were reacted with polyclonal rabbit antisera raised against purified H. pylori UreB and H. felis UreB MBP-fused proteins (sera diluted 1 : 5000). The difference in gel mobility of the respective non-recombinant UreB sub-units of H. felis and H. pylori can be seen. The numbers on the left refer to the molecular weights of standard marker proteins.

Figure 17 :

SDS-PAGE analysis of material eluted from the amylose column (lanes 2 and 3) or from the Ni-NTA column following elution : with buffer E (pH 4.5), lanes 4 and 5 ; or buffer C (pH 6.3), lanes 6 and 7. Material eluted from a lysate of MC1061 (PILL933) (lanes 2, 3, 5 and 7) and material eluted from a lysate of MC1061 (PMAL-c2) (lanes 4 and 6). Lane 3 contains the same material as in lane 2 except that it was resuspended in buffer E, thus demonstrating that buffer E is responsible for dimer formation of the MBP-HspA subunit, as seen in lanes 3 and 5.

Figure 18 :

Serum IgG responses to MBP (bottom), MBP-HspA (top) and MBP-HspB (middle) of 28 H. pylori infected patients (squares, left) and 12 uninfected patients (circles, right). The optical density of each serum in the ELISA assay described in Experimental procedures was read at 492 nm, after a 30 mn incubation. The sizes of the symbols are proportional to the number of sera giving the same optical density value.

EXAMPLES

I - CLONING, EXPRESSION AND SEQUENCING OF H. FELIS UREASE GENE :

EXPERIMENTAL PROCEDURES FOR PART I :

Bacterial strains and culture conditions :

H. felis (ATCC 49179) was grown on blood agar base no. 2 (Oxoid) supplemented with 5 % (v/v) lysed horse blood (BioMerieux) and an antibiotic supplement consisting of 10 ng ml^{-1} vancomycin (Lederle Laboratories), 2.5 μg ml^{-1} polymyxin B (Pfizer), 5μg ml^{-1} trimethoprim (Sigma Chemical Co.) and 2.5 μg ml^{-1} amphotericin B (E.R Squibb and Sons, Inc.). Bacteria were cultured on freshly prepared agar plates and incubated, lid uppermost, under microaerobic conditions at 37°C for 2-3 days. E. coli strains HB101 (Boyer and Roulland-Dussoix, 1969) and MC1061 (Maniatis et al., 1983), used in the cloning experiments, were grown routinely in Luria broth without glucose added or on Luria agar medium, at 37°C. Bacteria grown under nitrogen-limiting

conditions were passaged on a nitrogen-limiting solid medium consisting of ammonium-free M9 minimal medium (pH 7.4) supplemented with 0.4 % (w/v) D-glucose and 10 mM L-arginine
(Cussac et al., 1992).

DNA manipulations :

All standard DNA manipulations and analyses, unless mentioned otherwise, were performed according to the procedures described by Maniatis et al. (1983).

Isolation of H. felis DNA :

Total genomic DNA was extracted by an sarkosyl-proteinase K lysis procedure (Labigne-Roussel et al., 1988). Twelve blood agar plates inoculated with *H. felis* were incubated in an anaerobic jar (BBL) with an anaerobic gaspak (BBL 70304) without catalyst, for 1-2 days at 37°C. The plates were harvested in 50 ml of a 15 % (v/v) glycerol - 9 % (w/v) sucrose solution and centrifuged at 5,000 rpm (in a Sorvall centrifuge), for 30 min at 4°C. The pellet was resuspended in 0.2 ml 50 mM D-glucose in 25 mM Tris-10 mM EDTA (pH 8.0) containing 5 mg ml^{-1} lysozyme and transferred to a VTi65 polyallomer quick seal tube. A 0.2 ml aliquot of 20 mg ml^{-1} proteinase K and 0.02 ml of 5M sodium perchlorate were added to the suspension. Cells were lysed by adding 0.65 ml of 0.5M EDTA -10 % (w/v) Sarkosyl, and incubated at 65°C until the suspension cleared (approximately 5 min). The volume of the tube was completed with a CsCl solution consisting (per 100 ml) of 126 g CsCl, 1 ml aprotinine, 99 ml TES buffer (30 mM Tris, 5 mM EDTA, 50 mM NaCl (pH 7.5). Lysates were centrifuged at 45 000 rpm, for 15-18 h at 18°C. Total DNA was collected and dialysed against TE buffer (10 mM Tris, 1 mM EDTA), at 4°C.

Cosmid cloning :

Chromosomal DNA from H. felis was cloned into cosmid vector pILL575, as previoulsy described (Labigne et al, 1991). Briefly, DNA fragments arising from a partial digestion with Sau3A were sized on a (10 to 40 %) sucrose density gradient and then ligated into a BamHI-digested and dephosphorylated pILL575 DNA preparation. Cosmids were packaged into phage lambda particles (Amersham, In Vitro packaging kit) and used to infect E. coli HB101. To screen for urease expression, kanamycin-resistant transductants were replica-plated onto solid nitrogen-mimiting medium (see above) containing (20 $\mu g\ ml^{-1}$) kanamycin that had been dispensed into individual wells of microtitre plates (Becton Dickinson). The mictrotitre plates were incubated aerobically, at 37°C for 2 days before adding 0.1 ml urease reagent (Hazell et al., 1987) to each of the wells. Ureolysis was detected within 5-6 h at 37°C by a colour change in the reagent. Several urease-positive cosmid clones were restriction mapped and one was selected for subcloning.

Subcloning of H. felis DNA :

A large-scale CsCl plasmid preparation of cosmid DNA was partially digested Sau3A. DNA fragments (7 - 11 kb) were electroeluted from an agarose gel and purified using phenol-chloroform extractions. Following precipitation in cold ethanol, the fragments were ligated into Bg/III-digested plasmid pILL570 (Labigne et al., 1991) and the recombinant plasmids used to transform competent E. coli MC1061 cells. Spectinomycin-resistant transformants were selected and screened for urease expression under nitrogen-rich (Luria agar) and nitrogen-limiting conditions.

Quantitative ureas activity :

Cultures grown aerobically for 2.5 days at 37°C were harvested and washed twice in 0.85 % (w/v) NaCl. Pellets were resuspended in PEB buffer (0.1 M sodium phosphate buffer (pH 7.4) containing 0.01 M EDTA) and then sonicated by four 30-sec bursts using a Branson Sonifier model 450 set at 30 W, 50 % cycle. Cell debris was removed from the sonicates by centrifugation. Urease activities of the sonicates were measured in a 0.05 M urea solution prepared in PEB by a modification of the Berthelot reaction (Cussac et al., 1992). Urease activity was expressed as μmol urea $min^{-1}mg^{-1}$ bacterial protein.

Protein determination :

Protein concentrations were estimated with a commercial version of the bradford assay (Sigma Chemicals).

Transposon mutagenesis :

Random insertional mutations were generated within cloned H. felis via a MiniTn3-Km delivery system (Labigne et al., 1992). In brief, E. coli HB101 cells containing the transposase-encoding plasmid pTCA were transformed with plasmid pILL570 containing cloned H. felis DNA. Transposition of the MiniTn3-Km element into the pILL570 derivative plasmids was effected via conjugation. The resulting cointegrates were then selected for resolved structures in the presence of high concentrations of kanamycin (500 mgl-1) and spectinomycin (300 mgl-1).

SDS-PACE and Western blotting :

Solubilised cell extracts were analysed on slab gels, comprising a 4.5 % acrylamide stacking gel and 12.5 % resolving gel, according to the procedur of

Laemmli (Laemmli, 1970). Electrophoresis was performed at 200V on a mini-slab gel apparatus (Bio-Rad).

Proteins were transferred to nitrocellulose paper (Towbin et al., 1979) in a Mini Trans-Blot transfer cell (Bio-Rad) set at 100 V for 1 h (with cooling). Nitrocellulose membranes were blocked with 5 % (w/v) purified casein (BDH) in phosphate-buffered saline (PBS, pH 7.4) at room temperature, for 2 h (Ferrero et al., 1992). Membranes were reacted at 4°C overnight with antisera diluted in 1 % (w/v) casein prepared in PBS. Immunoreactants were then detected using a biotinylated secondary antibody (Kirkegaard and Perry Lab.) in conbination with avidin-peroxidase (KPL). A substrate solution composed of 0.3 % (w/v) 4-chloro-1-naphthol (Bio-rad) was used to visualise reaction products.

DNA Sequencing :

DNA fragments to be sequenced were cloned into M13mp18 and M13mp19 (Meissing and Vieira, 1982) bacteriophage vectors (Pharmacia). Competent E. coli JM101 cells were transfected with recombinant phage DNA and plated on media containing X-gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside) and isopropyl-β-D-thiogalactopyranoside. Plaques arising from bacteria infected with recombinant phage DNA were selected for the preparation of single-stranded DNA templates by polyethylene glycol treatment (Sanger et al., 1977). Single-stranted DNA sequenced according to the dideoxynucleotide chain termination method using a Sequenase kit (United States Biochemical Corp.).

Nucleotide sequence accession number :

The nucleotide accession number is X69080 (EMBL Data Library).

RESULTS OF PART I EXPERIMENTS :

Expression of urease activity by H. felis cosmid clones :

Cloning of partially digested fragments (30 to 45 kb in size) of H. felis chromosomal DNA into the cosmid vector pILL575 resulted in the isolation of approximately 700 cosmid clones. The clones were subcultured on nitrogen-limiting medium in order to induce urease expression (Cussac et al., 1992). Six of these were identified as being urease-positive after 5-6 h incubation (as described in the Experimental procedures section). No other urease-positive cosmid clones were identified, even after a further overnight incubation. Restriction enzyme analysis of 3 clones harbouring the urease-encoding cosmids revealed a common 28 kd DNA fragment. A cosmid (designated pILL199) containing DNA regions at both extremities of the common fragment was selected for subcloning.

Identification of H. felis genes required for urease expression when cloned in E. coli cells :

To define the minimum DNA region necessary for urease expression in E. coli cells, the urease-encoding cosmid pILL199 was partially digested with Sau3A and the fragments were subcloned into plasmid pILL570. The transformants were subcultured on nitrogen-rich and nitrogen-limiting media and screened for an urease-positive phenotype. Five transformants expressed urease activity when grown under nitrogen-limiting conditions, whereas no activity was detected following growth on nitrogen-rich medium. Restriction mapping analyses indicated that the urease-encoding plasmids contained inserts of between 7 and 11 kb. The

plasmid designated pILL205 was chosen for further studies.

Random mutagenesis of cloned H. felis DNA was performed to investigate putative regions essential for urease expression in E. coli and to localise the region of cloned DNA that contained the structural urease genes. Random insertion mutants of the prototype plasmid pILL205 were thus generated using the MiniTn3-Km element (Labigne et al, 1992). The site of insertion was restriction mapped for each of the mutated copies of pILL205 and cells harbouring these plasmids were assessed qualitatively for urease activity (figure 1). A selection of E. coli HB101 cells harbouring the mutated derivatives of pILL205 (designated "a" to "i") were then used both for quantitative urease activity determinations, as well as for the detection of the putative urease subunits by Western blotting.

The urease activity of E. coli HB101 cells harbouring pILL205 was 1.2 ± 0.5 μmol urea $min^{-1}mg^{-1}$ bacterial protein (table 1), which is approximately a fifth that of the parent H. felis strain used for the cloning. Insertion of the transposon at sites "a", "c", "d", "f" and "g" resulted in a negative phenotype, whilst mutations at sites "b", "e", "h" and "i" had no significant effect on the urease activities of clones harbouring these mutated copies of pILL205 (table 1). Thus mutagenesis of pILL205 with the MiniTn3-Km element identified three domains as being required for H. felis urease gene expression in E. coli cells.

Localisation of the H. felis urease structural genes :

Western blot analysis of extracts of E. coli cells harbouring pILL205 indicated the presence of two

polypeptides of approximately 30 and 66 kDa which cross-reacted with polyclonal H. felis rabbit antiserum (Figure 2A). These proteins were not produced by bacteria carrying the vector (pILL570). Native H. felis urease has been reported to be composed of repeating monomeric subunits with calculated molecular weights of 30 and 69 kDa (Turbett et al, 1992). Thus the 30 and 66 kDa proteins were thought to correspond to the ure A and ure B gene products, respectively. Interestingly an extract of E. coli cells harbouring the recombinant plasmid pILL763 (Cussac et al, 1992) containing the Helicobacter pylori ure A and ure B genes, expressed two polypeptides with approximate molecular sizes of 30 and 62 kDa which cross-reacted with the anti-H. felis antisera (figure 2B).

Table 1. Mutagenesis of *E. coli* clones and effect on urease activity.

plasmids [a]	Urease activity [b] (µmol urea min^{-1} mg^{-1} protein)
pILL205	1.2 ± 0.46 [c]
pILL205 :: a	neg [d]
pILL205 :: b	0.74 ± 0.32
pILL205 :: c	neg
pILL205 :: d	neg
pILL205 :: e	0.54 ± 0.15
pILL205 :: f	neg
pILL205 :: g	neg
pILL205 :: h	1.05 ± 0.25
pILL205 :: i	0.93 ± 0.35

[a] *E. coli* cells harboured pILL205 and its derivatives constructed by transposon mutagenesis. The letters correspond to the insertion sites of the MiniTn3-transposon on pILL205.

[b] Activities of bacteria grown aerobically for 3 days at 37 °C on solid M9 minimal medium supplemented with 10 mM L-arginine. The values represent the means ± standard deviations calculated from three determinations.

[c] Urease activity was approximately a fifth as large as that of *H. felis* wild-type strain (ATCC 49179) *i.e.* 5.7 ± 0.1 µmol urea min^{-1} mg^{-1} protein (Ferrero and Lee, 1991).

[d] No activity detected (limit of detection was < 1 nmol urea min^{-1} mg^{-1} of bacterial protein).

Clones harbouring the mutated derivatives of pILL205, in all but one case, expressed the ure A and ure B gene products (Figures 2A, B). Given that several of the mutants (i.e. mutants "c", "d", "f" and "g") synthesised the urease subunits yet did not produce an active enzyme, it is possible to speculate that accessory functions essential for urease activity may have been disrupted by transposon insertion. In contrast, the mutant designated pILL205::a did not produce the ure B product and was urease-negative. Thus the site of transposon insertion was presumed to be located in the ure B gene. Sequence analyses of the DNA region corresponding to insertion site "a" were undertaken to elucidate potential open reading frames encoding the structural polypeptides of H. felis urease.

Sequence analyses of H. felis structural urease genes :

Sequencing of a 2.4 kb region of H. felis DNA adjacent to transposon insertion site "a" resulted in the identification of two open reading frames (ORFs) designated ure A and ure B which are transcribed in the same direction (figure 3). The transposon was confirmed to be located at 240 bp upstream from the end of ure B. Both ORFs commenced with an ATG start codon and were preceded by a site similar to the E. coli consensus ribozome-binding sequence (Shine and Dalgarno, 1974). The intergenic space for the H. felis structural genes consisted of three codons which were in phase with the adjacent open-reading frames. This suggests that, as has already been observed to be the cas for Helicobacter pylori (Labigne et al, 1991), a single mutation in the stop codon of the ure A gene

would theoretically result in a fused single polypeptide.

The H. felis ure A and ure B genes encode polypeptides with calculated molecular weights of 26 074 kA and 61 663 Da, respectively, which are highly homologous at the amino-acid sequence level to the ure A and ure B gene products of H. pylori. The levels of identity between the corresponding ure A and ure B gene products of the two Helicobacter spp. was calculated to be 73.5 % and 88.2 % respectively. From the amino-acid sequence information, the predicted molecular weights of the ure A and ure B polypeptides from H. felis and H. pylori (Labigne et al, 1991) are very similar. Nevertheless the ure B product of H. felis had a lower mobility than the corresponding gene product from Helicobacter pylori when subjected to SDS-polyacrylamide gel electrophoresis (figure 2B)

II - EXPRESSION OF RECOMBINANT UREASE SUBUNIT PROTEINS FROM H. PYLORI AND H. FELIS : ASSESSMENT OF THESE PROTEINS AS POTENTIAL MUCOSAL IMMUNOGENS IN A MOUSE MODEL :

The aims of the study were to develop recombinant antigens derived from the urease subunits of H. pylori and H. felis, and to assess the immunoprotective efficacies of these antigens in the H. felis/mouse model. Each of the structural genes encoding the respective urease subunits from H. pylori and H. felis was independently cloned and over-expressed in Escherichia coli. The resulting recombinant urease antigens (which were fused to a 42 kDa maltose-binding protein of E. coli) were purified in large quantities from E. coli cultures and were immunogenic, yet enzymatically inactive. The findings demonstrated the

feasibility of developing a recombinant vaccine against H. pylori infection.

EXPERIMENTAL PROCEDURES FOR PART II :

Bacterial strains, plasmids and growth conditions :

H. felis (ATCC 49179) was grown on a blood agar medium containing blood agar base no. 2 (Oxoid) supplemented with 10% lysed horse blood (BioMérieux) and an antibiotic supplement consisting of vancomycin (10 μg/mL), polymyxin B (25 ng/mL), trimethoprim (5 μg/mL) and amphotericin B (2.5 μg/mL). Bacteria were cultured under microaerobic conditions at 37° C for 2 days, as described previously. E. coli strains MC1061 and JM101, used in cloning and expression experiments, were grown routinely at 37° C in Luria medium, with or without agar added. The antibiotics carbenicillin (100 μg/mL) and spectinomycin (100 μg/mL) were added as required.

DNA manipulations and analysis :

All DNA manipulations and analyses, unless mentioned otherwise, were performed according to standard procedures. Restriction and modification enzymes were purchased from Amersham (France). DNA fragments to be cloned were electroeluted from agarose gels and then purified by passage on Elutip mini-columns (Schleicher and Schull, Germany). Single-stranded DNA sequencing was performed using M13mp18 and M13mp19 bacteriophage vectors (Pharmacia, France). Single-stranded DNA templates were prepared from recombinant phage DNA by polyethylene glycol treatment. Sequencing of the templates was achieved according to the dideoxynucleotide chain termination

method using a Sequenase kit (United States Biochemical Corp., U.S.A.).

Preparation of inserts for cloning using the polymerase chain reaction (PCR) :

To clone the ureA genes of H. pylori and H. felis, degenerated 36-mer primers were conceived from the published urease sequences (Labigne et al., 1991 ; Ferrero and Labigne, 1993) (primer set #1 ; refer to table 2). Purified DNA from E. coli clones harbouring plasmids pILL763 and pILL207 (table 3), that encoded the structural genes of H. pylori and H. felis ureases, were used as template material in PCR reactions. Reaction samples contained : 10 - 50 ng of denatured DNA ; PCR buffer (50 mmol/L KCl in 10 mmol/L Tris-HCl [pH 8.3)]) ; dATP, dGTP, dCTP and dTTP (each at a final concentration of 1.25 mmol/L) ; 2.5 mmol/L $MgCl_2$; 25 pmol of each primer and 0.5 µL Taq polymerase. The samples were subjected to 30 cycles of the following programme : 2 min at 94° C, 1 min at 40° C.

The amplification products were cloned into the cohesive ends of the pAMP vector (figure 1) according to the protocol described by the manufacturer ("CloneAmp System", Gibco BRL ; Cergy Pontoise, France). Briefly, 60 ng of amplification product was directly mixed in a buffer (consisting of 50 mmol/L KCl, 1.5 mmol/L $MgCl_2$, 0.1 % (wt/vol) gelatine in 10 mmol/L Tris-HCl, pH 8.3) with 50 ng of the pAMP 1 vector DNA and 1 unit of uracil DNA glycolsylase. Ligation was performed for 30 min at 37° C. Competent cells (200 µL) of E. coli MC1061 were transformed with 20 µL of the ligation mixture. Inserts were subsequently excised from the polylinker of the pAMP vector by double digestion with BamH1 and Pst1, and

then subcloned into the expression vector pMAL (New England Biolabs Inc., Beverly, USA) chosen for the production of recombinant antigens (pILL919 and pILL920, respectively, figure 13), as well as in M13mp bacteriophage for sequencing.

Amplification of a product containing the ureB gene of H. pylori was obtained by PCR using a couple of 35-mer primers (set #2, table 2). The PCR reaction mixtures were first denatured for 3 min at 94° C, then subjected to 30 cycles of the following programme : 1 min at 94° C, 1 min at 55° C and 2 min at 72° C. The purified amplification product (1850 bp was digested with EcoRI and PstI and then cloned into pMAL (pILL927, figure 2). Competent cells of E. coli MC1061 were transformed with the ligation reaction.

H. felis ureB was cloned in a two-step procedure, that allowed the production of both complete and truncated versions of the UreB subunit. Plasmid pILL213 (table 3) was digested with the enzymes DraI, corresponding to amino acid residue number 219 of the UreB subunit and HindIII. The resulting 1350 bp fragment was purified and cloned into pMAL that had been digested with XmnI and HindIII (pILL219, figure 2). In order to produce a clone capable of synthesizing a complete UreB protein, PCR primers were developed (set #3, table 2) that amplified a 685 bp fragment from the N-terminal portion of the ureB gene (excluding the ATG codon), that also overlapped the beginning of the insert in plasmid pILL219. The PCR amplified material was purified and digested with bamHI and HindIII, and then cloned into pMAL (pILL221, figure 14). A 1350 bp PstI-PstI fragment encoding the remaining portion of the UreB gene product was subsequently excised from pILL219 and cloned into a

linearised preparation of pILL221 (pILL222, figure 14).

<u>Expression of recombinant urease polypeptides in the vector pMAL :</u>

The expression vector pMAL is under the control of an inducible promoter (P_{lac}) and contains an open-reading frame (ORF) that encodes the production of MalE (Maltose-binding protein, MBP). Sequences cloned in-phase with the latter ORF resulted in the synthesis of MBP-fused proteins which were easily purified on amylose resin. Of the two versions of pMAL that are commercially available, the version not encoding a signal sequence (ie. pMAL-c2) synthesized greater amounts of recombinant proteins and was thus used throughout.

<u>E. coli</u> clones harbouring recombinant plasmids were screened for the production of fusion proteins, prior to performing large-scale purification experiments.

<u>Purification of recombinant urease polypeptides :</u>

Fresh 500 mL volumes of Luria broth, containing carbenicillin (100 µg/mL and 2% (wt/vol) glucose, were inoculated with overnight cultures (5 mL) of <u>E. coli</u> clones. The cultures were incubated at 37° C and shaken at 250 rpm, until the A_{600} = 0.5. Prior to adding 1 mmol/L (final concentration) isopropyl-β-D-thiogalactopyranoside (IPTG) to cultures, a 1.0 mL sample was taken (non-induced cells). Cultures were incubated for a further 4 h at which time another 1.0 mL sample (induced cells) was taken. The non-induced and induced cell samples were later analysed by SDS-PAGE.

IPTG-induced cultures were centrifuged at 7000 rpm for 20 min, at 4° C and the supernatant discarded. Pellets were resuspended in 50 mL column buffer (200 mmol/L NaCl, 1 mmol/L EDTA in 10 mmol/L TrisHCl,pH 7.4), containing the following protease inhibitors (supplied by Boehringer, Mannheim, Germany) : 2 µmol/L leupeptin, 2 µmol/L pepstatin and 1 mmol/L phenylmethylsulphonyl fluoride (PMSF). Intact cells were lysed by passage through a French Pressure cell (16 000 lb/in^2). Cell debris was removed by centrifugation and lysates were diluted in column buffer to give a final concentration of 2.5 mg protein/mL, prior to chromatography on a 2.6 cm x 20 cm column of amylose resin (New England Biolabs). The resin was washed with column buffer at 0.5 mL/min until the A_{280} returned levels. The MBP-fused recombinant proteins were eluted from the column by washing with column buffer containing 10 mmol/L <u>l-maltose</u>.

Fractions containing the recombinant proteins were pooled and then dialysed several times at 4° C against a low salt buffer (containing 25 mmol/L NaCl in 20 mmol/L TrisHCl, pH 8.0). The pooled fractions were then loaded at a flow rate of 0.5 mL/min onto a 1.6 x 10 cm anion exchange column (HP-Sepharose , Pharmacia, Sweden) connected to a Hi-Load chromatography system (Pharmacia). Proteins were eluted from the column using a salt gradient (25 mmol/L to 500 mmol/L NaCl). Fractions giving high absorbance readings at A_{280} were exhaustively dialysed against distilled water at 4° C and analysed by SDS-PAGE.

<u>Rabbit antisera :</u>

Polyclonal rabbit antisera was prepared against total cell extracts of H. pylori strain 85P (Labigne et al., 1991) and H. felis (ATCC49179). Polyclonal rabbit antisera against recombinant protein preparations of H. pylori and H. felis urease subunits was produced by immunizing rabbits with 100 µg of purified recombinant protein in Freund's complete adjuvant (Sigma). Four weeks later, rabbits were booster-immunized with 100 µg protein in Freund's incomplete adjuvant. On week 6, the animals were terminally bled and the sera kept at -20° C.

Protein analyzes by SDS-PAGE and western blotting :

Solubilized cell extracts were analyzed on slab gels, comprising a 4.5% acrylamide stacking gel and a 10% resolving gel, according to the procedure of Laemmli. Electrophoresis was performed at 200 V on a mini-slab gel apparatus (Bio-Rad, USA).

Proteins were transferred to nitrocellulose paper in a Mini Trans-Blot transfer cell (Bio-Rad) set at 100 V for 1 h, with cooling. Nitrocellulose membranes were blocked with 5% (wt/vol) casein (BDH, England) in phosphate-buffered saline (PBS, pH 7.4) with gentle shaking at room temperature, for 2 h. Membranes were reacted at 4° C overnight with antisera diluted in 1% casein prepared in PBS. Immunoreactants were detected using specific biotinylated seondary antibodies and streptavidin-peroxidase conjugate (kirkegaard and Parry Lab., Gaithersburg, USA). Reaction products were visualized on autoradiographic film (Hyperfilm, Amersham, France) using a chemiluminescence technique (ECL system, Amersham).

Protein concentrations were determined by the Bradford assay (Sigma Chemicals corp., St Louis, USA).

Animal xperimentation :

Six week old female Swiss Specific Pathogen-Free (SPF) mice were obtained (Centre d'Elevage R. Janvier, Le-Genest-St-Isle, France) and maintained on a commercial pellet diet with water ad libitum. The intestines of the animals were screened for the absence of Helicobacter muridarum. For all orogastric administrations, 100 µL aliquots were delivered to mice using 1.0 mL disposable syringes, to which polyethylene catheters (Biotrol, Paris, France) were attached.

Preparation of sonicated extracts and inocula from H. felis cultures :

H. felis bacteria were harvested in PBS and centrifuged at 5000 rpm, for 10 min in a Sorvall RC-5 centrifuge (Sorvall, USA) at 4° C. The pellets were washed twice and resuspended in PBS. Bacterial suspensions were sonicated as previously described and were subjected to at least one freeze-thaw cycle. Protein determinations were carried out on the sonicates.

To ensure a virulent culture of H. felis for protection studies, H. felis bacteria were maintained in vivo until required. Briefly, mice were inoculated three times (with 10^{10} bacteria/mL), over a period of 5 days. The bacteria were reisolated from stomach biopsies on blood agar medium (4 - 7 days' incubation in a microaerobic atmosphere at 37° C). Bacteria grown for two days on blood agar plates were harvested directly in peptone water (Difco, USA). Bacterial viability and motility was assessed by phase microscopy prior to administration to animals.

Mous protecti n studi s :

Fifty μg of recombinant antigen and 10 μg cholera holotoxin (Sigma Chemical Corp.), both resuspended in HCO_3, were administrated orogastrically to mice on weeks 0, 1, 2 and 3. Mice immunized with sonicated H. felis extracts (containing 400 - 800 μg of total protein) were also given 10 μg of cholera toxin. On week 5, half of the mice from each group were challenged with an inoculum of virulent H. felis. The remainder of the mice received an additional "boost" immunization on week 15. On week 17 the latter were challenged with a culture of H. felis.

Assessment of H. felis colonisation of the mouse :

Two weeks after receiving the challenge dose (ie. weeks 7 and 19, respectively) mice were sacrificed by spinal dislocation. The Stomachs were washed twice in sterile 0.8% NaCl and a portion of the gastric antrum from each stomach was placed on the surfaces of 12 cm x 12 cm agar plates containing a urea indicator medium (2% urea, 120 mg Na_2HPO_4, 80 mg KH_2PO_4, 1.2 mg phenol red, 1.5 g agar prepared in 100 mL). The remainder of each stomach was placed in formal-saline and stored until processed for histology. Longitudinal sections (4μm) of the stomachs were cut and routinely stained by the Giemsa technique. When necessary, sections were additionally stained by the Haematoxylin-Eosin and Warthin-Starry silver stain techniques;

The presence of H. felis bacteria in mouse gastric mucosa was assessed by the detection of urease activity (for up to 24 h) on the indicator medium, as well as by the screening of Giemsa-stained gastric sections that had been coded so as to eliminate observer bias. The numbers of bacteria in gastric sections were semi-quantitatively scored according to the following scheme : 0, no bacteria seen throughout

sections ; 1, few bacteria (< 20) seen throughout ; 2, occasional high power (H.P.) field with low numbers (< 20) of bacteria ; 3, occasional H.P. field with low to moderate numbers (< 50) of bacteria ; and 4, numerous (> 5) H.P. fields with high numbers of bacteria (> 50). Mononuclear cell infiltrates were scored as follows : 0, no significant infiltration ; 1, infiltration of low numbers of mononuclear cells limited to the submucosa and muscularis mucosa ; 2, infiltration of moderate numbers of mononuclear cells to the submucosa and muscularis mucosa, sometimes forming loose aggregates ; and 3, infiltration of large numbers of mononuclear cells and featuring nodular agglomerations of cells.

RESULTS OF PART II EXPERIMENTS :

Expression of Helicobacter urease polypeptides in E. coli :

Fragments containing the sequences encoding the respective UreA gene products of H. felis and H. pylori were amplified by PCR and cloned in-phase with an ORF encoding the 42 kDa MBP, present on the expression vector pMAL. Sequencing of the PCR products revealed minor nucleotidic changes that did not, however, alter the deduced amino acid sequences of the respective gene products. E. coli MC1061 cells transformed with these recombinant plasmids (pILL919 and pILL920, respectively) expressed fusion proteins with predicted molecular weights of approximately 68 kDa. Following chromatography on affinity (amylose resin) and anion exchange gel media (Q-Sepharose), these proteins were purified to high degrees of purity (figure 1). The yield from 2-L cultures of recombinant

E. coli cells was approximately 40 mg of purified antigen.

Similarly, the large UreB subunits of H. pylori and H. felis ureases were expressed in E. coli (plasmids pILL927 and pILL222, respectively) and produced fusion proteins with predicted molecular weights of 103 kDa. The yield in these cases was appreciably lower than for the UreA preparations (approximately 20 mg was recovered from 2-L of bacterial culture). Moreover, problems associated with the cleavage of the UreB polypeptides from the MBP portion of the fusion proteins were encountered. These difficulties were attributed to the large sizes of the recombinant UreB polypeptides.

Analysis of the recombinant urease polypeptides :

Western blot analyses of the antigen preparations with rabbit polyclonal antisera raised to whole-extracts of H. pylori and H. felis bacteria demonstrated that the antigens retained immunogenicity to the homologous as well as heterologous antisera (figures 14 and 15). The antisera did not recognize the MBP component alone. Cross-reactivity between the urease polypeptides of H. pylori and H. felis was consistent with the high degrees of identity between the amino acid sequences of these proteins.

Rabbit polyclonal antisera raised against purified recombinant UreA and UreB proteins prepared from H. pylori and H. felis strongly reacted with the urease polypeptides present in whole-cell extracts of the bacteria (figure 16). As we had already observed, the UreB subunit of H. felis urease migrated slightly higher on SDS-PAGE gels than did that of H. pylori (figure 16).

Pr paration of H. felis inocula used in immunoprotection studies :

To ensure the virulence of H. felis bacterial inocula, bactera were reisolated from H. felis-infected mouse stomachs (see Materials and methods). The bacteria were passaged a minimum number of times in vitro. Stock cultures prepared from these bacteria, and stored at -80° C, were used to prepare fresh inocula for other mouse protection studies. This procedure ensured that the inocula used in successive experiments were reproducible.

Immunization of mice against gastric H. felis infection :

Mice that had been immunized for three weeks with the given antigen preparations were divided into two lots and one half of these were challenged two weeks later with an H. fellis inoculum containing 10^7 bacteria/mL. One group of animals that had been immunized with recombinant H. felis UreA were also challenged but, unlike the other animals, were not sacrificed until week 19.

a) Protection at week 5 :

Eighty-five % of stomach biopsy samples from the control group of mice immunized with H. felis sonicate preparations were urease-negative and therefore appeared to have been protected from H. felis infection (table 4). This compared to 20% of those from the other control group of animals given MBP alone. The proportion of urease-negative stomachs for those groups of mice given the recombinant urease subunits varied from 70% (for H. pylori UreB) to 20% (for H. pylori UreA).

The levels of bacterial colonisation by H. felis was also assessed from coded histological slides prepared from gastric tissue. Due to the striking helical morphology of H. felis bacteria, the organisms could be readily seen on the mucosal surfaces of both gastric pit and glandular regions of the stomach. Histological evidence indicated that the levels of protection in mice was lower than that observed by the biopsy urease test : 25% and 20% of gastric tissue from mice immunized with H. felis sonicate preparations of H. pylori UreB, respectively, were free of H. felis bacteria.

Amongst certain groups of these mice the preponderance of urease-negative biopsies, as well as lower histological scores for bacterial colonisation (unpublished data), suggested that an immunoprotective response had been elicited in the animals. This response, however, may have been insufficient to protect against the inoculum administered during the challenge procedure.

b) Protection at week 17 :

The remaining mice, from each group of animals, were boosted on week 15. These mice were challenged at week 17 with an H. felis inoculum containing approximately 100-fold less bacteria than that used previously. Two weeks later all stomach biopsies from the MBP-immunized mice were urease-positive (table 4). In contrast, urease activity for gastric biopsies from mice immunized with the recombinant urease subunits varied from 50% for H. pylori UreA to 100% for H. felis UreB. The latter was comparable to the level of protection observed for the group of animals immunized with H. felis sonicated extracts. Histological evidence demonstrated that the UreB subunits of H.

felis and H. pylori protected 60% and 25% of immunized animals, respectively. This compared with a level of 85% protection for mice immunized with H. felis sonicated extracts. Immunization of mice with recombinant H. pylori UreA did not protect the animals. Similarly, the stomachs of all H. felis UreA-immunized mice, that had been challenged at week 5, were heavily colonised with H. felis bacteria at week 19 (table 4).

The urease gastric biopsy test, when compared to histological analysis of gastric tissue sections, gave sensitivity and specificity values of 63% and 95%, respectively. Thus histology proved to be the more accurate predictor of H. felis infection in the mouse.

Cellular immune response in immunized stomachs :

In addition to the histological assessment of H. felis colonisation, mouse gastric tissue was also scored (from 0 to 3) for the presence of a mononuclear cell response. In mice immunized with MBP alone, a mild chronic gastritis was seen with small numbers of mononuclear cells restricted to the muscularis mucosa and to the submucosa of the gastric epithelium. In contrast, there were considerable numbers of mononuclear cells present in the gastric mucosae from animals immunized with either the recombinant urease polypeptides, or with H. felis sonicate preparations. These inflammatory cells coalesced to form either loose aggregates, in the submucosal regions of the tissue, or nodular structures that extended into the mucosal regions of the gastric epithelia. The mononuclear cell response did not appear to be related to the presence of bacteria as the gastric mucosae from the H. felis UreA-immunized mice, that were

heavily colonized with H. felis bacteria, contained little or no mononuclear cells.

Table 2 The oligomeric primers used in PCR-based amplification of urease-encoding nucleotide sequences.

Primer set		Nucleotide sequence (5' -> 3')
# 1	forw	...CAU CCT* AAA^{G} GAA^{G} T^{C}TA* GAT^{C} AAA^{G} T^{C}TA* ATG
	rev	T^{C}TC C^{T}TT A*CG A*CG A*G^{C}A^{T} A^{G,T}AT C^{T}TT C^{T}TT CAT CUA...
#2	forw	CC GGA GAA TTC ATT AGC AGA AAA GAA TAT GTT TCT ATG *EcoRI*¥ (GAA TTC)
	rev	AC GTT CTG CAG CTT ACG AAT AAC TTT TGT TGC TTG AGC *PstI*¥ (CTG CAG)
#3	forw	GGA TCC AAA AAG ATT TCA CG *BamHI*¥ (GGA TCC)
	rev	GGA AGC TT C TGC AGG TGT GCT TCC CCA GTC *HindIII*¥ (A AGC TT) *PstI*¥ (C TGC AG)

* Degenerated nucleotides in which all possible permutations of the genetic code were included (A, T, G, C).

G,C,T The given nucleotides were degenerated with the specific base(s) shown.

¥ Restriction sites introduced in the amplified fragments.

Table 3 Plasmids used

Plasmid	Vector	Relevant phenotype or character	Reference
pILL763	pILL570	9.5 kb fragment (*Sau*3a partial digest of *H. pylori* chromosome) (Sp^R)	Cussac *et al.*, 1991
pILL199	pILL575	35 kb fragment (*Sau*3A partial digest of *H. felis* chromosome)	Ferrero & Labigne,'93
pILL207	pILL570	11 kb fragment (*Sau*3A partial digest of pILL199)	This study
pILL919	pMAL-C2	0.8 kb *Bam*HI-*Pst*I [a] insert containing a nucleotide fragment encoding H. felis ureA gene (Ap^R)	This study
pILL920	pMAL-C2	0.8 kb *Bam*HI-*Pst*I[a] insert containing PCR product encoding *H. pylori ure*A gene	This study
pILL927	pMAL-C2	1.8 kb *Eco*RI-*Pst*I[a] PCR fragment encoding *H. pylori ure*B gene	This study
pILL213	pUC19	2 kb fragment resulting from *Sau*3A partial digest of pILL207 (Ap^R)	This study
pILL219	pMAL-C2	1.4 kb *Dra*I-*Hind*III[b] insert containing *H. felis ure*B (bases 657 - 1707)	This study
pILL 221	pMAL-C2	0.7 kb *Bam*HI-*Pst*I PCR fragment encoding *H. felis ure*B (bases 4 - 667)	This study
pILL222	pMAL-C2	1.35 kb *Pst*I-*Pst*I[c] fragment encoding *H. felis ure*B (bases 667 - 1707) from pILL219 cloned into linerized pILL221	This study

Table 4 Protection of mice by immunization with recombinant urease proteins.

Antigen	Protection (%) [a]	
	Urease	Histology
MBP	0 % (0/10)	0 % (0/10)
UreA *H. pylori*	50 (4/8)	0 (0/10)
UreA *H. felis* [b]	12.5 (1/8)	0 (0/10)
UreB *H. pylori*	65 (5/8)	25 (2/8)
UreB *H. felis*	100 (7/7)	60 (5/7)
H. felis sonicate	100 (8/8)	85 (7/8)

a Challenge inoculum dose was 10^5 bacteria/mouse

b Mice were challenged on week 5 (with 10^7 bacteria) and were sacrificed on week 19.

III- HELICOBACTER PYLORI hspA-B HEAT SHOCK GENE CLUSTER : NUCLEOTIDE SEQUENCE, EXPRESSION AND FUNCTION :

A homolog of the heat shock proteins (HSPs) of the GroEL class, reported to be closely associated with the urease of Helicobacter pylori (a nickel metalloenzyme), has recently been purified from H. pylori cells by Dunn et al, and Evans et al. (Infect. Immun. 60:1946, 1992, 1946 and 2125, respectively). Based on the reported N-terminal amino acid sequence of this immunodominant protein, degenerated oligonucleotides were synthesized in order to target the gene (hspB) encoding the GroEL-like protein in the chromosome of H. pylori strain 85P. Following gene amplification, a 108-base pair (bp)-fragment encoding the 36 first amino acids of the HspB protein was purified, and used a probe to identify in the H. pylori genomic bank a recombinant cosmid harboring the entire HspB encoding gene. The hspB gene was mapped to a 3.15 kilobases (kb) BglII restriction fragment of the pILL684 cosmid. The nucleotide sequence of that fragment subcloned into the pILL570 plasmid vector (pILL689) revealed the presence of two open reading frames (OFRs) designated hspA and hspB, the organization of which was very similar to be groESL bicistronic operons of other bacterial species. hspA and hspB encode polypeptides of 118 and 545 amino acids respectively, corresponding to calculated molecular masses of 13.0 and 58.2 kilodaltons (kDa), respectively. Amino acid sequence comparison studies revealed i) that the H. pylori HspA and HspB protein were highly similar to their bacterial homologs; ii) that the HspA H. pylori protein features a striking motif at the carboxyl terminus that other bacterial

GroEs-homologs lack ; this unique motif consists of a series of eight histidine residues resembling metal binding domain, such a nickel binding. Surprisingly, immediately upstream of the gene cluster an IS5 insertion element was found that was absent in the H. pylori genome, and was positively selectionned during the cosmid cloning process. The IS5 was found to be involved in the expression of the hspA and hspB genes in pILL689. The expression of the HspA and HspB proteins from the pILL689 plasmid was analyzed in minicell-producing strain. Both polypeptides were shown to be constitutively expressed in the E. coli cells. When the pILL689 recombinant plasmid was introduced together with the H. pylori urease gene cluster into an E. coli host strain, an increase of urease activity was observed suggesting a close interaction between the heat shock proteins and the urease enzyme. Supporting the concept of a specific function for the HspA chaperone, was the fact that whereas a single hspB copy was found in the H. pylori genome, two copies of the hspA were found in the genome, one linked to the hspB gene and one unlinked to the hspB gene. Attempts to construct isogenic mutants of H. pylori in the hspA and the hspB gene were unsucesseful suggesting that these genes are essential for the survival of the bacteria.

EXPERIMENTAL PROCEDURES FOR PART III :

Bacterial strains, plasmids, and culture conditions :

The cloning experiments were performed with genomic DNA prepared from H. pylori strain 85P. H. pylori strain N6 was used as the recipient strain for the electroporation experiments because of its favorable transformability. E. coli strain HB101 or

strain MC1061 were used as a host for cosmid cl ning and subcloning experiments, respectively. E. coli P678-54 was used for preparation of minicells. Vectors and recombinant plasmids used in this study are listed in Table 1. H. pylori strains were grown on horse blood agar plates, supplemented with vancomycin (10 mg/l), polymyxin B (2,500 U/I), trimethoprim (5 mg/l), and amphotericin B (4 mg/l). Plates were incubated at 37°C under microaerobic conditions in an anaerobic jar with a carbon dioxide generator envelope (BBL 70304). E. coli strains were grown in L-broth without glucose (10 g of tryptone, 5 g of yeast extract, and 5 g of NaCI per liter ; pH 7.0) or on L-agar plates (1.5 % agar) at 37°C. For measurement of urease activity, the nitrogen-limiting medium used consisted of ammonium-free M9 minimal agar medium (pH7.4) containing 0.4 % D-glucose as the carbon source, and freshly prepared filter-sterilized L-arginine added to the final concentration of 10 mM. Antibiotic concentrations for the selection of recombinant clones were as follows (in milligrams per liter) : kanamycin, 20 ; spectinomycin, 100 ; carbenicillin, 100.

Preparation of DNA :

Genomic DNA from H. pylori was prepared as previously described. Cosmid and plasmid DNAs were prepared by an alkaline lysis procedure followed by purification in cesium chloride-ethidium bromide gradients as previously described.

Cosmid cloning :

The construction of the cosmid gene bank of H. pylori 85P in E. coli HB101, which was used for the cloning of the H. pylori hspA-B gene cluster, has been described previ usly.

DNA analysis and cloning meth dology :

Restriction endonucleases, T4 DNA ligase, DNA polymerase I large (Klenow) fragment, and Taq polymerase were purchased from Amersham, T4 DNA polymerase from Biolabs, and calf intestinal phosphatase from Pharmacia. All enzymes were used according to the instructions of the manufacturers. DNA fragments were separated on agarose gels run in Tris-acetate buffer. The 1-kb ladder from Bethesda Research Laboratories was used as a fragment size standard. When necessary, DNA fragments were isolated by electroelution from agarose gels as previously described and recovered from the migration buffer by means of an Elutip-d minicolumn (Schleicher and Schuell, Dassel, Germany). Basic DNA manipulations were performed according to the protocols described by Sambrook et al.

Hybridization :

Colony blots for screening of the H. pylori cosmid bank and for identification of subclones were prepared on nitrocellulose membranes (Schleicher and Schuell, Dassel, Germany) according to the protocol of Sambrook et al. (43). Radioactive labelling of PCR-products was performed by random priming, using as primers the random hexamers from Pharmacia. Colony hybridizations were performed under high stringency conditions (5 x SSC, 0.1 % SDS, 50 % formamide, 42° C) (1 x SSC ; 150 mM NaCl, 15 mM sodium citrate, pH 7.0). For Southern blot hybridizations, DNA fragments were transferred from agarose gels to nitrocellulose sheets (0.45-μm pore size ; Schleicher & Schuell, Inc.), and hybridized under low stringency conditions (5 x SSC, 0.1 % SDS, 30 or 40 % formamide, at 42° C with ^{32}P-labeled deoxyribonucleotide pr bes Hybridization was

revealed by autoradiography using Amersham Hyperfilm-MP.

DNA sequencing :

Appropriate fragments of plasmid DNA were subcloned into M13 mp 18/19 vectors. Single stranded DNA was prepared by phage infection of E. coli strain JM101. Sequencing was performed by the dideoxynucleotide chain termination method using the United States Biochemicals Sequenase kit. Both the M13 universal primer and additional specific primers (Fig.1) were used to sequence both the coding and non-coding DNA strands. Sequencing of double-stranded DNA was performed as previously described. Direct sequencing of PCR product was carried out following purification of the amplified, electroeluted PCR product through an Elutip-d minicolumn (Schleicher & Schuell) ; The classical protocol for sequencing using the Sequenase kit was then used with the following modifications : PCR product was denatured by boiling annealing mixture containing 200 picomoles of the oligonucleotide used as primer and DMSO to the final concentration of 1 % for 3 minutes ; the mixture was then immediatly cool on ice ; the labeling step was performed in presence of manganese ions (mM).

Electroporation of H. pylori :

In the attempt to construct H. pylori mutants, appropriate plasmid constructions carrying the targeted gene disrupted by a cassette containing a kanamycin resistance gene (aph3'-III), were transformed into H. pylori strain N6 by means of electroporation as previously described. Plasmid pSUS10 harboring the kanamycin disrupted flaA gene was used as positive control of electroporation. After

electroporation, bacteria were grown on non-selective plates for a period of 48 h in order to allow for the expression of the antibiotic resistance and then transferred onto kanamycin-containing plates. The selective plates were incubated for up to 6 days.

Polymerase chain reaction (PCR) :

PCRs were carried out using a Perkin-Elmer Cetus thermal cycler using the GeneAmp kit (Perkin-Elmer Cetus). Classical amplification reaction involved 50 picomoles (pmoles) of each primer and at least 5 pmoles of the target DNA. The target DNA was heat denatured prior addition to the amplification reaction. Reaction consisted of 25 cycles of the following three steps : denaturation (94° C for 1 minute), annealing (at temperatures ranging between 42 and 55° C, depending on the calculated melting temperatures of the primers, for 2 min), and extension (72° C for 2 min). When degenerated oligonucleotides were used in non stringent conditions, up to 1000 pmoles of each oligonucleotide were added, 50 cycles were carried out, and annealing was performed at 42° C.

Analysis of proteins expressed in minicells :

Minicells harboring the appropriate hybrid plasmid were isolated and labeled with [^{35}S] methionine (50 μ Ci/ml). Approximately 100,000 cpm of acetone-precipitable material was subjected to sodium dodecyl sulfate (SDS) -polyacrylamide gel electrophoresis in a 12.5 % gel. Standard proteins with molecular weights ranging from 94,000 to 14,000 (low< molecular-weights kit from Bio-Rad Laboratories) were run in parallel. The gel was stained and examined by fluorography, using En^{3}Hance (New England Nuclear).

Urease activity :

Urease activity was quantitated by the Berthelot reaction by using a modification of the procedure which has already been described. Urease activity was expressed as micromoles of urea hydrolyzed per minute per milligram of bacterial protein.

RESULTS OF PART III EXPERIMENTS :

Identification of a recombinant cosmid harboring the Helicobacter pylori GroEL-like heat shock protein encoding gene :

Based on the published N-terminal amino sequence of the purified heat shock protein of H. pylori, two degenerated oligonucleotides were synthesized to target the gene of interest in the chromosome of H. pylori strain 85P. The first one 5' - **G C N A A R G A R A T H A A R T T Y T C N G** - 3' where **N** stands for the four nucleotides, **R** = A and G, **Y** = T and C, **H** = T, C, and A, is derived from for the first 8 amino acids of the protein (AKEIKFSD) ; the second one 5' - **C R T T N C K N C C N C K N G G N C C C A T** - 3', where **K** = G and T, corresponds to the complementary codons specifying the amino acid from position 29 to position 36 (MGPRGRNV, ref). The expected size for the PCR product was 108 base pairs (bp). The amplification reaction was performed under low stringency conditions as described in the "Materials and Methods" section, and led to the synthesis of six fragments with size ranging from 400 bp to 100 bp. The three smallest fragments were electroeluted from an acrylamide gel, and purified. Direct sequencing of the PCR products permitted the identification of a DNA fragment encoding an amino acid sequence corresponding to the published sequence. This fragment was therefore

labeled and used as probe in colony hybridization to identify recombinant cosmids exhibiting homology to a 5' segment of the H. pylori GroEL-like encoding gene ; this gene was further designated hspB. The gene bank consists of 400 independent kanamycin-resistant E. coli transductants harboring recombinant cosmids. Of those one single clone hybridized with the probe, and harbored a recombinant plasmid designated pILL684, 46 kb in size. The low frequency observed when detecting the hspB gene (1 of 400) was unusual when compared with that of several cloned genes which were consistently detected in five to seven recombinant cosmids. In order to identify the hspB gene, fragments with sizes of 3 to 4 kb were generated by partial restriction of the pILL684 cosmid DNA with endonuclease Sau3A, purified, and ligated into the BglII site of plasmid vector pILL570. Of 100 subclones, x were positive clones, and one was further studied (pILL689) ; it contains a 3.15 kb insert, flanked by two BglII restriction sites, that was mapped in detail (Fig. 5). Using the PCR ^{32}P labeled probe, the 5' end of the hspB gene was found to map to the 632 bp HindIII-SphI central restriction fragment of pILL689, indicating that one could expect the presence of the entire hspB gene in the pILL689 recombinant plasmid.

DNA sequence and deduced amino acid sequence of the H. pylori hspA-B gene cluster :

The 3200 bp of pILL689 depicted in Fig. 5 were sequenced by cloning into M13mp18 and M13mp19, the asymetric restriction fragments BglII-SphI, SphI-HindIII, HindIII-BglII ; each cloned fragment was independently sequenced on both strands 16 oligonucleotide primers (Fig.1) were synthesized to

confirm the reading and/or to generate sequences overlapping the independently sequenced fragments ; these were used as primers in double-stranded-DNA sequencing analyses.

The analysis of the sequence revealed two distinct genetic elements. First the presence of two open reading frames (ORFs), depicted in figure 5, transcribed in the same direction, that were designated hspA and hspB ; The nucleotide sequence and the deduced amino acid sequence of the two ORFs are presented in Fig. 6. The first codon of hspA begins 323 bp upstream of the leftward HindIII site of pILL689 (Fig. 5) and is preceded by a Shine-Dalgarno ribosome-binding site (RBS) (**GGAGAA**). The hspA ORF codes for a polypeptide of 118 amino acids. The initiation codon for the hspB ORF begins 25 nucleotides downstream the hspA stop codon ; it is preceded by a RBS site (**AAGGA**). The hspB ORF encodes a polypeptide of 545 amino acids and is terminated by a TAA codon followed by a palindromic sequence resembling a rho-independent transcription terminator (free energy, ΔG = -19.8 kcal/mol) (Fig. 6). The N-terminal amino acid sequence of the deduced protein HspB was identical to the N-terminal sequence of the purified H.pylori heat shock protein previously published with the exception of the N-terminal methionine, which is absent from the purified protein and might be posttranslationally removed, resulting in a mature protein of 544 amino acids.

The deduced amino acid sequences of H. pylori HspA and HspB were compared to several amino acid sequences of HSPs of the GroES and GroEL class (Fig. 7). HspB exhibited high homology at the amino acid level with the Legionella pneumophila HtpB protein (82.9 % of similarities), with the Escherichia coli

GroEL protein (81.0 % of similarities), with the Chlamydia psittaci or C. trachomatis HypB protein (79.4 % of similarities), with Clostridium perfringens Hsp60 protein (80.7 % of similarities), and to a lesser extent to the GroEL-like proteins of Mycobacterium. However, like almost all the GroEL homologs, H. pylori HspB demonstrated the conserved carboxyl-terminus glycine-methionine motif (MGGMGGMGGMGGMM) which was recently shown to be dispensable in the E. coli GroEL chaperonin. The degree of homology at the amino acid level between the H. pylori HspA protein and the other GroES-like proteins is shown in Fig. 7. The alignment shown features a striking motif at the carboxyl terminus of the H. pylori HspA protein that other bacterial GroES-homologs lack. This unique highly charged motif consists of 27 additional amino acids capable of forming a loop between two double cystein residues ; ot the 27 amino acids, 8 are histidine residues highly reminiscent of a metal binding domain.

The second genetic element revealed by the sequence analysis, was the presence of an insertion sequence (IS5) 84 bp upstream of the hspA gene. The nucleotide sequence of this element matched perfectly that previously described for IS5 in E. coli, with the presence of a 16 nucleotide sequence (CTTGTTCGCACCTTCC) that corresponds to one of the two inverted repeats which flank the IS5 element. Because of the perfect match at the DNA level, we suspected that the IS5 was not initially present in the H. pylori chromosome, but had rather inserted upstream of the hspA-HspB gene cluster during the cloning process, a hypothesis that needed to be confirmed by further analyses.

Identificati n of the upstream sequence of th hspA-B gene cluster in H. pylori chromosome :

The presence of the IS5 was examined by gene amplification using two oligonucleotides, one being internal to the IS5 element and the other one downstream of the IS5 element (oligo #1 and #2, Fig. 6), to target a putative sequence i) in the chromosome of H. pylori strain 85P, ii) in the initial cosmid pILL684, and iii) in the 100 subclones resulting of the Sau3A partial restriction of the pILL684 recombinant cosmid. IS5 was absent from the chromosome of H. pylori, and was present in the very first subcultures of the E. coli strain harboring cosmid pILL684. Among the 100 pILL684 subclone derivatives which appeared to contain all or part of the IS5 sequence, we then looked for a subclone harboring the left end side of the IS5 plus the original upstream sequence of the hspA-hspB gene cluster. This screening was made by restriction analysis of the different Sau3A partial generated subclones. The restriction map of one (pILL694) of the plasmids fulfilling these criteria is shown in Fig. 5. The left end side of the IS5 nucleotide sequence was determined ; the presence of a 4-bp duplication CTAA on both side of the 16-bp inverted repeats of the IS5 element (Fig. 6) allowed us to confirm the recent acquisition of the IS5 element by transposition. A 245-nucleotide sequence was then determined that mapped immediately upstream of the IS5 element (shown Fig. 6). This sequence consists of a non coding region in which the presence of a putative consensus heat shock promoter sequence was detected ; it shows a perfectly conserved -35 region (**TAACTCGCTTGAA**) and a less consentaneous -10 region (CTCAATTA). Two oligonucleotides (#3 and #4, shown on Fig.2) were synthesized which mapped to

sequences located on both side of the IS5 element present in the recombinant cosmid ; these two oligonucleotides should lead to the amplification of a XXXXbp fragment when the IS5 sequence is present and a fragment in the absence of the IS5. The results of the PCR reaction using as target DNA the pILL684 cosmid, the pILL694 plasmid, and the H. pylori 85P chromosome fit the predictions (results not shown). Moreover, direct sequencing of the PCR product obtained from the H. pylori chromosome was performed and confirmed the upstream hspA-hspB reconstructed sequence shown in Fig. 6 (B). To further confirm the genetic organization of the whole sequenced region, two probes were prepared by gene amplification of the pILL689 plasmid using oligonucleotides #5 and #6, and #7 and #8 (Fig. 6). ; they were used as probes in Southern hybridization experiments under low stringency conditions against an HindIII digest of the H. pylori 85P chromosme. The results demonstrate that no other detectable rearrangement had occured during the cloning process (data not shown). These experiments allowed us to demonstrate that whereas a single copy of the hspB gene was present in the chromosome of H. pylori strain 85, two copies of the hspA gene were detected by Southern hybridization.

Analysis of polypeptides expressed in minicells :

The pILL689 and the pILL692 recombinant plasmids and the respective cloning vectors pILL570, and pACYC177, were introduced by transformation into E. coli P678-54, a minicell-producing strain. The pILL689 and pILL692 plasmids (Fig. 5) contain the same 3.15-kb insert cloned into the two vectors. pILL570 contains upstream of the poly-cloning site a stop of transcription and of translati n ; the orientati n of

the insert in pILL689, was made in such way that the transcriptinnal stop was located upstream of the IS5 fragment and therefore upstream of the hspA and HspB genes. Two polypeptides that migrated with polypeptides having apparent molecular weights of 60 kDa and 14 kDa were clearly detected in minicell-experiments from pILL689 and pILL692 (results not shown), whereas they were absent from the corresponding vectors ; these results indicated that the hspA and hspB genes were constitutively expressed from a promoter located within the IS5 were constitutively expressed from a promoter located within the IS5 element. Moreover, whereas the amount of polypeptides visualized on the SDS gel was in good agreement with the copy number of the respective vectors, the intensity of the two polypeptidic bands suggested a polycistronic transcription of the two genes.

Attempts to understand the role of the Hspa and HspB proteins :

Two disruptions of genes were achieved in E. coli by inserting the Km cassette previously described within the hspA or the hspB gene of plasmids pILL686 and pILL691. This was done in order to return the disrupted genes in H. pylori by electroporation, and to select for allelic replacement. The pILL696 resulting plasmid encoded a truncated form of the HspA protein, corresponding to the deletion of the C-terminal end amino acid sequence ; in that plasmid the Km cassette was inserted in such way that the promoter of the Km gene could serve as promoter for the hspB downstream gene. The pILL687 and pILL688 plasmids resulted from the insertion of the Km cassette in either orientation within the hspB gene. None of these

constructs led to the isolation of kanamycin transformants of H. pylori strain N6, when purified pILL687, pILL688, pILL696 plasmids (Table 2, Fig. 5) were used in electroporation experiments, whereas the pSUS10 plasmid used as positive control always did. These results suggest the H. pylori HspA and HspB protein are essential proteins for the survival of H. pylori.

Because of i) the constant description in the literature of a close association of the HspB protein with the urease subunits ; -ii) the unique structure of the HspA protein with the C-terminal sequence reminiscent of a nickel binging domain, and iii) of the absence of viable hspA and/or hspB mutants of H. pylori, we attempted to demonstrate a role of the H. pylori Hsps proteins in relations with the H. pylori urease by functional complementation experiments in E. coli. Plasmids pILL763 or pILL753 (both pILL570 derivatives, Table 5) encoding the urease gene cluster were introduced with the compatible pILL692 plasmid (pACYC177 derivative) that constitutively expresses the HspA et HspB polypeptides as visualized in minicells. In both complementations, the expression of the HspA and HspB proteins in the same E. coli cell allows to observe a three fold increase in the urease activity following induction of the urease genes on minimum medium supplemented with 10 mM L- Arginine as limiting nitrogen source.

Table 5 : Vectors and hybrid plasmids used in this study.

Plasmid	Vector	Size (kb)	Characteristics (a)	Origin or Reference
	pILL575	10	Mob, Cos, Km	-
	pILL570	5.3	Mob, Sp	-
	pACYC177	3.9	Ap,Km	-
pILL600	pBR322	5.7	Ap, Km, source of Km-cassette	-
pILL684	pILL575	46	Mob, Km, cosmid containing *H. pylori hspA-B*	*Sau3A* partial digest of *H. pylori* 85P DNA
pILL685	pILL570	9.29	Mob, Sp, plasmid containing *H. pylori hspB*	*Sau3A* partial digest of pILL684
pILL686	pUC19*c	4.5	Ap, plasmid containing *H. pylori hspB*	1.9-kb *BglII-ClaI* pILL685 cloned into pUC19*
	pUC19*(c)	5.9	Ap, Km, *H. pylori hspB* Ω Km-orientation A(b)	1.4-kb *SmaI-SmaI* pILL600 cloned into pILL686
pILL688	pUC19*(c)	5.9	Ap, Km, *H. pylori hspB* Ω Km- orientation B (b)	1.4-kb *SmaI-SmaI* pILL600 cloned into pILL686
pILL689	pILL570	8.45	Mob, Sp, plasmid containing *H. pylori hspA-B*	*Sau3A* partial digest of pILL684
pILL691	pUC19**(c)	3.9	Ap, plasmid containing *H.pylori hspA* 1.3-kb	*SphI-SphI* pILL689 cloned into pUC19**
pILL692	pACYC177	7.05	Ap, Km, plasmid containing *H. pylori hspA-B*	3.15-kb*BglII* pILL689 cloned into pACYC177
pILL694	pILL570	8.7	Sp, plasmid containing left end of IS5	*Sau3A* partial digest of pILL684
pILL696	pUC19**(c)	5.3	Ap, Km, *H. pylori hspA* Ω Km-orientation A (b)	1.4-kb *SmaI-SmaI* pILL600 cloned into pILL691
pSUS10	pIC20R2	7.7	Ap, Km,*H. pylori flaA* Ω Km	-
pILL753	pILL570	16.5	Sp, plasmid containing *ureA,B,C,D,E,F,G,H,I*	-
pILL763	pILL570	14.75	Sp, plasmid containing *ureA,B,E,F,G,H,I* -	-

(a) Mob, conjugative plasmid due to the presence of OriT; Ap, Km, and Sp, resistance to ampicillin, kanamycin, and spectinomycin, respectively; Cos, presence of lambda cos site.
(b) Orientation A indicates that the Kanamycin promoter initiates transcription in the same orientation as that of the gene where the cassette has been inserted; orientation B, the opposite.
(c) pUC19* ane pUC19** : derivatives from pUC19 vector in which the the *SphI* and *HindIII* site, respectively, have been end-filled by using the Klenow polymerase and self religated.

IV - EXPRESSION, PURIFICATION AND IMMUNOGENIC PROPERTIES OF H. PYLORI HSPA AND HSPB :

EXPERIMENTAL PROCEDURE FOR PART IV :

Expression and purification of recombinant fusion proteins :

The MalE-HspA, and MalE-HspB fusion proteins were expressed following the cloning of the two genes within the pMAL-c2 vector as described in the "Results" section using the following primers :

oligo #1 ccggagaattcAAGTTTCAACCATTAGGAGAAAGGGTC
oligo #2 acgttctgcagTTTAGTGTTTTTTGTGATCATGACAGC
oligo #3 ccggagaattcGCAAAAGAAATCAAATTTTCAGATAGC
oligo #4 acgttctgcagATGATACCAAAAAGCAAGGGGGCTTAC

Two liters of Luria medium containing glucose (30%) and ampicillin (100 μg/ml) were inoculated with 20 ml of an overnight culture of strain MC1061 containing the fusion plasmid and incubated with shaking at 37°C. When the OD600 of the culture reached 0.5, IPTG (at a final concentration of 10 mM) was added, and the cells were incubated for a further 4 hours. Cells were harvested by centrifugation (5000 rpm for 30 min at 4°C), resuspended in 100 ml of column buffer consisting of 10 mM Tris-HCl, 200 mM NaCl, 1 mM EDTA supplemented with protease inhibitors [(Leupeptin (2μM) - Pepstatin (2μm) - PMSF (1mM) - Aprotinin (1:1000 dilution)], and passed through a French press. After centrifugation (10,000 rpm for 20 min at 4°C), the supernatant were recovered and diluted (2-fold) with column buffer. The lysate was filtered through a 0.2 μm nitrocellulose filter prior to loading onto a preequilibrated amylose resin (22 x 2.5 cm). The fusion proteins were eluted with a 10mM maltose solution prepared in column buffer, and the fractions

containing the fusion proteins were pooled, dialyzed against distilled water, and lyophilized. Fusion proteins were resuspended in distilled water at a final concentration of 2 mg of lyophilized material/ml, and stored at -20°C. Concentration and purity of the preparations were controlled by the Bradford protein assay (Sigma Chemicals) and SDS-PAGE analyses.

Nickel binding properties of recombinant proteins :

E. coli MC1061 cells, containing either the pMAL-c2 vector or derivative recombinant plasmids, were grown in 100 ml-Luria broth in the presence of carbenicillin (100 μg/ml). The expression of the genes was induced with IPTG for four hours. The cells were centrifuged and the pellet was resuspended in 2 ml of Buffer A (6M guanidine hydrochloride, 0.1 M NaH_2PO_4, 0.01Tris, pH8.0). After gentle stirring for one hour at room temperature, the suspensions were centrifuged at 10,000 g for 15 min at 4°C. A 1.6 ml aliquot of Nickel-Nitrilo-Tri-Acetic resin (Nickel-NTA, QIA express), previously equilibrated in Buffer A, was added to the supernatant and this mixture was stirred at room temperature for one hour prior to loading onto a column. The column was washed with 20 ml buffer A, then 30 ml buffer B (8M urea, 0.1M Na-phosphate, 0.01MTris-HCl, pH8.0). The proteins were eluted successively with the same buffer as buffer B adjusted to pH 6.3 (Buffer C), pH 5.9 (Buffer D) and pH 4.5 (Buffer E) and Buffer F (6M guanidine hydrochloride, 02M acetic acid). Fifty μl of each fraction were mixed with 50 μl of SDS buffer and loaded on SDS gels.

Human sera :

Serum samples were obtained from 40 individuals, 28 were H. pylori-infected patients as confirmed by a positive culture for H. pylori and histological examination of the biopsy, and 12 were uninfected patients. The sera were kindly provided by R. J. Adamek (University of Bochum, Germany).

Immunoblotting :

Upon completion of SDS-PAGE runs in a Mini-PROTEAN II electrophoresis cell, proteins were transferred to nitrocellulose paper in a Mini Trans-Blot transfer cell (Bio-Rad) set at 100 V for 1 h (with cooling). Immunostaining was performed as previously described (Ferrero et al., 1992), except that the ECL Western blotting detection system (Amersham) was used to visualize reaction products . Human sera and the rabbit antiserum, raised against a whole-cell extract of H. pylori strain 85P, were diluted 1:1000 and 1:5000, respectively, in 1% (w/v) casein prepared in phosphate-buffered saline (PBS, pH7.4).

Serological methods [enzyme-linked immunosorbent assay, (ELISA)] :

The following quantities of antigens were absorbed onto 96-well plates (Falcon 3072) : 2.5 μg of protein MalE, 5 μg of MalE-HspA, or 2.5 μg of MalE-HspB. The plates were left overnight at 4°C, then washed 3 times with ELISA wash solution (EWS) [1% PBS containing 0.05% (v/v) Tween 20]. Saturation was achieved by incubating the plates for 90 min at 37°C in EWS supplemented with 1% milk powder. Wells were again washed 3 times with EWS and then gently agitated for 90 min at 37°C in the presence of human sera (diluted 1:500 in EWS with 0.5% milk powder), under

agitation. Bound imunoglobulins were detected by incubation for 90 min at 37°C with biotinylated secondary antibody (goat anti-human IgG, IgA or IgM diluted [1:1000] in EWS supplemented with 0.5% milk powder) in combination with streptavidin-peroxidase (1:500) (Kirkegaard and Perry Lab.). Bound peroxidase was detected by reaction with the citrate substrate and hydrogen peroxide. Plates were incubated in the dark, at room temperature, and the optical density at 492 nm was read at intervals of 5, 15 and 30 min in an ELISA plate reader. After 30 min, the reaction was stopped by the addition of hydrochloric acid to a final concentration of 0.5M.

RESULTS OF PART IV EXPERIMENTS :

Construction of recombinant plasmids producing inducible MalE-HspA, and HspB fusion proteins :

The oligonucleotides #1 and #2 (hspA) and #3 and #4 (hspB) were used to amplify by PCR the entire hspA and the hspB genes, respectively. The PCR products were electroeluted, purified and restricted with EcoRI and PstI. The restricted fragments (360 bp and 1600 bp in size, respectively) were then ligated into the EcoRI-PstI restricted pMAL-c2 vector to generate plasmids designated pILL933 and pILL934, respectively. Following induction with IPTG, and purification of the soluble protein on amylose columns, fusion proteins of the expected size (55 kDa for pILL933 [figure 17], and 100 kDa for pILL9334) were visualized on SDS-PAGE gels. Each of these corresponded to the fusion of the MalE protein (42.7 kDa) with the second amino-acid of each of the Hsp polypeptides. The yield of the expression of the fusion proteins was 100 mg for

MalE-HspA and 20 mg for MalE-HspB when prepared from 2 liters of broth culture.

Study of the antigenicity of the HspA and HspB fusion proteins, and of the immunogenicity of HspA and HspB in patients infected with H. pylori :

In order to determine whether the fusion proteins were still antigenic, each was analyzed by Western blot with rabbit antiserum raised against the MalE protein and a whole-cell extract of H. pylori strain 85P. Both fusion proteins were immunoreactive with antibody to MalE (not shown) and with the anti-H. pylori antiserum. The anti-H. pylori antiserum did not recognize the purified MalE protein (figure 18). These results demonstrated that the fusion proteins retained their antigenic properties ; in addition, whereas the HspB protein was known to be immunogenic, this is the first demonstration that HspA per se is immunogenic in rabbits.

In the same way, in order to determine whether the HspA and HspB polypeptides were immunogenic in humans, the humoral immune response against HspA and/or HspB in patients infected with H. pylori was analyzed and compared to that of uninfected persons using Western immunoblotting assays and enzyme-linked immunosorbent assays (ELISA). None of the 12 sera of the H. pylori-negative persons gave a positive immunoblot signal with MalE, MalE-HspA, or MalE-HspB proteins (figure 18). In contrast, of 28 sera from H. pylori-positive patients, 12 (42.8%) reacted with the HspA protein whilst 20 (71.4%) recognized the HspB protein. All of the sera that recognized HspA also reacted with the HspB protein. No association was observed between the immune response and the clinical presentation of the H. pylori infection alth ugh such

a conclusion might be premature because of the small number of strains analyzed.

<u>**Nickel binding properties of the fused MalE-HspA protein :**</u>

MBP-HspA recombinant protein expressed following induction with IPTG, was purified from a whole cell extract by one step purification on nickel affinity column whereas the MBP alone, nor MBP-HspB exhibited this property. Figure 18 illustrates the one step purification of the MBP-HspA protein that was eluted as a monomer at pH6.3, and as a monomer at pH4.5. The unique band seen in panel 7 and the two bands seen in panel 5 were both specifically recognized with anti-HspA rabbit sera. This suggested that the nickel binding property of the fused MBP-HspA protein might be attributed to the C-terminal sequence os HspA which is rich in Histidine and Cysteine residues.

References

Boyer, H. W., and Roulland-Dussoix, D (1969) A complementation analysis of the restriction and modification of DNA in Escherichia coli. J Mol Biol 41: 459-472.

Chen, M., Lee, A., and Hazell, S. L. (1992) Immunisation against gastric helicobacter infection in a mouse/Helicobacter felis model.Lancet 339: 1120-1121.

Corthesy-Theulaz, I. et al (1993), Acta Gastro-Enterol. Belgica Suppl., vol. 56, p 64 (VIth Workshop on Gastroduodenal pathology and H. pylori).

Cover, T. L., Puryear, W.; Perez-Perez, G.J., and Blaser, M. (1991) Effect of urease on HeLa cell vacuolation induced by Helicobacter pylori cytotoxin.Infect Immun 59: 1264-1270.

Cussac, V., Ferrero, R. L., and Labigne, A. (1992) Expression of Helicobacter pylori urease genes in Escherichia coli grown under nitrogen-limiting conditions.J Bacteriol 174: 2466-2473.

Davin, C. et al., Abstract A-304, Gastroenterology 1993 (Abstract supplement).

Dick-Hegedus, E., and Lee, A. (1991) The use of a mouse model to examine anti-Helicobacter pylori agents. Scand J Gastroenterol 26: 909-915.

Dick E., Lee A., Watson G., and O'Rourke J. (1989) Use of the mouse for the isolation and investigation of stomach-associated, spiral-helical shaped bacteria from man and other animals. J Med Microbiol 29: 55-62.

Dunn, B. E., R.M., Roop II, C.-C. Sung, S.A. Sharma, G.I. Perez-Perez, and M.J. Blaser, 1992. Identification and purification of a cpn60 heat

sh ck protein homolog from Helicobacter pylori. Infect Immun. 60: 1946-1951.

Eaton, K. A., Brooks, C. L., Morgan, D. R., and Krakowka, S. (1991) Essential role of urease in pathogenesis of gastritis induced by Helicobacter pylori in gnotobiotic piglets. Infect Immun 59: 2470-2475.

Evans, D. J., Evans, D. G., Engstrand, L. and Graham, D. Y. (1992) heat shock protein of Helicobacter pylori. Infect Immun 60: 2125-2127.

Ferrero, R. L., and Lee, A. (1991) The importance of urease in acid protection for the gastric-colonising bacteria Helicobacter pylori and Helicobacter felis sp. nov. Microb Ecol Hlth Dis 4:121-134.

Ferrero, R. L., Cussac, V., Courcoux, P. and Labigne, A. (1992) Construction of isogenic urease-negative mutants of Helicobacter pylori by allelic exchange. J Bacteriol 174:4212-4217.

Ferrero, R. L. and Labigne, A. (1993) Molec. Microbiol. 9, 323-333.

Freedburg, A. S., and Barron, L. E. (1940) The presence of spirochetes in human gastric mucosa. American Journal of Digestive Diseases 7:443-445.

Goodwin, C. S., Zrmstrong, J. A., Chilvers, T., Peters, M., Collins, M. D., Sly, L., McConnell, W., and Harper, W. E. S.(1989) Transfer of Campylobacter pylori comb. nov. and Helicobacter mustelae comb. nov., respectively. Int J Syst Bacteriol 39: 397-405.

Hazell, S. L., and Lee, A. (1986) Campylobacter pyloridis,urease, hydrogen ion back diffusion, and gastric ulcers. Lancet ii: 15-17.

Hazell, S. L., Borody, T. J., Gal, A., and Lee, A. (1987) Campylobacter pyloridis gastritis I:

Detection of urease as a marker of bacterial colonization and gastritis. Am J Gastroenterol 82: 292-296.

Hu, L-T, Foxall, P. A., Russell, R., and Mobley, H. L. T. (1992) Purification of recombinant Helicobacter pylori urease apoenzyme encoded by ureA and ureB. InfectImmun. 60:2657-2666.

Jones, B. D., and Mobley, H. L. T. (1989) Proteus mirabilis urease : nucleotide sequence determination and comparison with jack bean urease. J Bacteriol 171:6414-6422.

Krakowka, S., Morgan D. R., Kraft W. G., and Leunk R. D.(1987) Establishment of gastric Campylobacter pylori infection in the neonatal gnotobiotic piglet. Infect Immun 55:2789-2796.

Labigne-Roussel, A., Courcoux, P., and Tompkins, L. (1988) Gene disruption and replacement as a feasible approach for mutagenesis of Campylobacter jejuni. J Bacteriol 170:1704-1708.

Labigne, A., Cessac, V., and Courcoux, P. (1991) Shuttle cloning and nucleotide sequences of Helicobacter pylori genes responsible for urease activity. J Bacteriol 173:1920-1931.

Labigne, A., Courcoux, P., and Tompkins, L. (1992) Cloning of Campylobacter jejuni genes required for leucine biosynthesis, and construction of leu-negative mutant of C. jejuni by shuttle transposon mutagenesis. Res Microb 143: 15-26.

Laemmli, E. K. (1970) Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227: 680-685.

Lee, A., Hazell, S. L., O'Rourke, J., and Kouprach, S. (1988) Isolation of a spiral-shaped bacterium from the cat stomach. Infect Immun 56: 2843-2850.

Lee, A., Fox, J. G., Otto, G., and Murphy, J. (1990) A small animal model of human Helicobacter pylori active chronic gastritis. Gastroenterol 99: 1315-1323.

Lee, M. H., Mulrooney, S. B., Renner, M. J., Marckowicz, Y., and Hausinger, R. P. (1992) Klebsiella aerogenes urease gene cluster: Sequence of ure D and demonstration that four accessory genes (ure D, ure E, ure F, and ure G) are involved in nikel metallocenter biosynthesis. J Bacteriol 174: 4324-4330.

Luger, A., and Neuberger, H. (1921) Uber spirochatenbefunde im magensaft und der diagnostische Bedeutung für das carcinoma ventriculi. Zeit Klin Med 92: 54.

Mai, U. E. H., Perez-Perez, G. I., Allen, J. B., Wahl, S. M., Blaser, M. J., and Smith, P. D. (1992) surface proteins from Helicobacter pylori exhibit chemotactic activity for human leukocytes and are present in gastric mucosa. J Exp Med 175: 517-525.

Maniatis, T., Fritsch, E., and Sambrook, J. (1983) Molecular cloning : a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor N. Y.

Marshall, B. J., Royce, H., Annear, D. I., Goodwin, C. D., Pearman, J. W., Warren, J. R., and Armstrong, J. A. (1984) Original isolation of Campylobacter pyloridis from human gastric mucosa. Microbios Lett 25: 83-88.

Marshall, B. J., Barrett, L. J., Prakash, C., McCallem, R. W., and Guerrant, R. L. (1990) Urea protects Helicobacter (Campylobacter) pylori from the bactericidal effect of acid. Gastroenterol 99: 697-702.

Meissing, J., and Vieira, J. (1982) A new pair of M13 vectors for selecting either DNA strand of double-digest restriction fragments. Gene 19: 269-276.

Mobley, H. L. T., and Hausinger, R. P. (1989) Microbial ureases: significance, regulation, and molecular characterisation. Microbiol Rev 53: 85-108.

Newell, D. G., Lee, A., Hawtin, P. R., Hudson, M. J., Stacey, A. R., and Fox, J. (1989) Antigenic conservation of the ureases of spiral-and helical-shaped bacteria colonising the stomachs of man and animals. FEMS Microbiol Lett 65:183-186.

Nomura, A., Stermmermann, G. N., Ghyou, P-H., Kato, I., Perez-Perez, G. I., and Blaser, M. J. (1991) Helicobacter pylori infection and gastric carcinoma among Japanese Americans in Hawaii. N Eng J Med 325: 1132-1136.

Parsonnet, J., Friedman, G. D., Vanderstee, D. P., Chang, Y., Vogelman, J. H., Orentreich, N., and R. Sibley (1991) Helicobacter pylori infection and the risk of gastric carcinoma. N Eng J Med 325: 1127-1131.

Paster, B. J., Lee, A., Dewhirst, F. E., Fox, J. G., Tordoff, L. A., Fraser, G. J., O'Rourke, J. L., Taylor, N. S., and Ferrero, R. (1990) The phylogeny of Helicobacter felis sp. nov., Helicobacter mustalae, and related bacteria. Int J Syst Bacteriol 41: 31-38.

Peterson, W. L. (1991) Helicobacter pylori and peptic ulcer disease. N Engl J Med 324: 1043-1047.

Radin, J. M., Eaton, K. A., Krakowka, S., Morgan, D. R., Lee, A., Otto, G., and Fox, J. G. (1990)

Helicobacter pylori infection in gnotobiotic dogs. Infect Immun 58: 2606-2612.

Salomon, H. (1896) Ueber das Spirillem des Saugetiermagens und sein Verhalten zu den Belegzellen. Zentral Bakteriol Parasiten Infektion 19: 433-442.

Sanger, F., Nicklen, S., and Coulson, A. R. (1977) DNA sequencing with chain terminating inhibitors. Proc Natl Acad Sci USA 74: 5463-5467.

Shine, J., and Dalgarno, L. (1974) The 3'-terminal sequence of Escherichia coli 16S ribosomal RNA: complementarity to nonsense triplets and ribosome binding sites. Proc Natl Acad Sci USA 71: 1342-1346.

Sidebotham, R. L., and Baron, J. H. (1990) Hypothesis : Helicobacter pylori, urease, mucus, and gastric ulcer. Lancet 335: 193-195.

Smoot, D. T., Mobley, H. L. T., Chippendale, G. R., Lewinson, J. F., and Resau, J. H. (1990) Helicobacter pylori urease activity is toxic to human gastric epithelial cells. Infect Immun 58: 1992-1994.

Solnick, J.V., et al, Infec. and Immunity, May 1994, p 1631-1638.

Towbin, H., Staehelin, T., and Gordon, J. (1979) Electrophoretic transfer of proteins from polyacrylamide gels to nitrocellulose sheets: procedure and some applications. Proc Natl Acad Sci 76: 4350-4354.

Turbett, G. R., Nandapalan, N., Campbell, I. G., Nikoletti, S. M., and Mee, B. J. (1991) Characterization of the urease from Helicobacter pylori and comparison with the ureases from related spiral gastric bacteria. FEMS Microbiol Immunol 76: 19-24.

Turbett, G. R., Hoj, P., Horne, R., and Mee, B. J. (1992) Purification and characterization of the urease enzymes of Helicobacter species from humans and animals. Infect Immun 60: 5259-5266.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT:
(A) NAME: INSTITUT PASTEUR
(B) STREET: 25-28 rue du Dr Roux
(C) CITY: PARIS CEDEX 15
(E) COUNTRY: FRANCE
(F) POSTAL CODE (ZIP): 75724
(G) TELEPHONE: 45.68.80.94
(H) TELEFAX: 40.61.30.17

(A) NAME: INSTITUT NATIONAL DE LA SANTE ET DE LA RECHERCHE MEDICALE
(B) STREET: 101 rue de Tolbiac
(C) CITY: PARIS CEDEX 13
(E) COUNTRY: FRANCE
(F) POSTAL CODE (ZIP): 75654
(G) TELEPHONE: 44.23.60.00
(H) TELEFAX: 45.85.07.66

(ii) TITLE OF INVENTION: IMMUNOGENIC COMPOSITIONS AGAINST HELICOBACTER INFECTION, POLYPEPTIDES FOR USE IN THE COMPOSITIONS AND NUCLEIC ACID SEQUENCES ENCODING SAID POLYPEPTIDES.

(iii) NUMBER OF SEQUENCES: 8

(iv) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)

(v) CURRENT APPLICATION DATA:
APPLICATION NUMBER: EP 93401309.5

(2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2619 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 31..36
(D) OTHER INFORMATION: /standard_name= "Shine-Dalgarno sequence"

(ix) FEATURE:
(A) NAME/KEY: misc_feature
(B) LOCATION: 756..759
(D) OTHER INFORMATION: /standard_name= "Shine-Dalgarno sequence"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 43..753
(D) OTHER INFORMATION: /standard_name= "URE A"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 766..2475
(D) OTHER INFORMATION: /standard_name= "URE B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

```
TGATAGCTTG GCTACCAATA GAAATTCAAT AAGGAGTTTA GG ATG AAA CTA ACG       54
                                              Met Lys Leu Thr
                                                1

CCT AAA GAA CTA GAC AAG TTA ATG CTC CAT TAT GCG GGC AGA TTG GCA     102
Pro Lys Glu Leu Asp Lys Leu Met Leu His Tyr Ala Gly Arg Leu Ala
  5                  10                  15                  20

GAA GAA CGC TTG GCG CGT GGT GTG AAA CTC AAT TAC ACC GAA GCG GTC     150
Glu Glu Arg Leu Ala Arg Gly Val Lys Leu Asn Tyr Thr Glu Ala Val
                 25                  30                  35

GCG CTC ATT AGC GGG CGT GTG ATG GAA AAG GCG CGT GAT GGT AAT AAA     198
Ala Leu Ile Ser Gly Arg Val Met Glu Lys Ala Arg Asp Gly Asn Lys
             40                  45                  50

AGC GTG GCG GAT TTG ATG CAA GAA GGC AGG ACT TGG CTT AAA AAA GAA     246
Ser Val Ala Asp Leu Met Gln Glu Gly Arg Thr Trp Leu Lys Lys Glu
         55                  60                  65

AAT GTG ATG GAC GGC GTA GCA AGC ATG ATT CAT GAA GTG GGG ATT GAA     294
Asn Val Met Asp Gly Val Ala Ser Met Ile His Glu Val Gly Ile Glu
     70                  75                  80

GCT AAC TTC CCC GAT GGA ACC AAG CTT GTA ACT ATC CAC ACT CCG GTA     342
Ala Asn Phe Pro Asp Gly Thr Lys Leu Val Thr Ile His Thr Pro Val
 85                  90                  95                 100

GAG GAT AAT GGC AAA TTA GCC CCC GGC GAG GTC TTC TTA AAA AAT GAG     390
Glu Asp Asn Gly Lys Leu Ala Pro Gly Glu Val Phe Leu Lys Asn Glu
                105                 110                 115

GAC ATT ACT ATT AAC GCC GGC AAA GAA GCC ATT AGC TTG AAA GTG AAA     438
Asp Ile Thr Ile Asn Ala Gly Lys Glu Ala Ile Ser Leu Lys Val Lys
            120                 125                 130
```

```
AAT AAA GGC GAT CGT CCT GTG CAG GTG GGA TCA CAT TTC CAC TTC TTC      486
Asn Lys Gly Asp Arg Pro Val Gln Val Gly Ser His Phe His Phe Phe
        135                 140                 145

GAA GTG AAT AAG CTC TTG GAC TTC GAT CGC GCA AAA AGC TTT TGC AAA      534
Glu Val Asn Lys Leu Leu Asp Phe Asp Arg Ala Lys Ser Phe Cys Lys
    150                 155                 160

CGC CTA GAC ATT GCA TCT GGA ACA GCG GTG CGC TTT GAA CCC GGG GAG      582
Arg Leu Asp Ile Ala Ser Gly Thr Ala Val Arg Phe Glu Pro Gly Glu
165                 170                 175                 180

GAA AAA AGT GTG GAA CTC ATT GAC ATC GGC GGG AAT AAG CGC ATC TAT      630
Glu Lys Ser Val Glu Leu Ile Asp Ile Gly Gly Asn Lys Arg Ile Tyr
                185                 190                 195

GGC TTT AAT TCT TTG GTG GAT CGC CAA GCC GAT GCC GAT GGT AAA AAA      678
Gly Phe Asn Ser Leu Val Asp Arg Gln Ala Asp Ala Asp Gly Lys Lys
            200                 205                 210

CTC GGC TTA AAA CGC GCT AAA GAA AAA GGT TTT GGG TCT GTA AAC TGC      726
Leu Gly Leu Lys Arg Ala Lys Glu Lys Gly Phe Gly Ser Val Asn Cys
        215                 220                 225

GGT TGT GAA GCG ACT AAA GAT AAA CAA TAAGGAAAAA CC ATG AAA AAG        774
Gly Cys Glu Ala Thr Lys Asp Lys Gln                Met Lys Lys
    230                 235                                 1

ATT TCA CGA AAA GAA TAT GTT TCT ATG TAT GGT CCC ACT ACC GGG GAT      822
Ile Ser Arg Lys Glu Tyr Val Ser Met Tyr Gly Pro Thr Thr Gly Asp
    5                   10                  15

CGT GTT AGA CTC GGC GAC ACT GAT TTG ATC TTA GAA GTG GAG CAT GAT      870
Arg Val Arg Leu Gly Asp Thr Asp Leu Ile Leu Glu Val Glu His Asp
20                  25                  30                  35

TGC ACC ACT TAT GGT GAA GAG ATC AAA TTT GGG GGC GGT AAA ACT ATC      918
Cys Thr Thr Tyr Gly Glu Glu Ile Lys Phe Gly Gly Gly Lys Thr Ile
                40                  45                  50

CGT GAT GGG ATG AGT CAA ACC AAT AGC CCT AGC TCT TAT GAA TTA GAT      966
Arg Asp Gly Met Ser Gln Thr Asn Ser Pro Ser Ser Tyr Glu Leu Asp
            55                  60                  65

TTG GTG CTC ACT AAC GCC CTC ATT GTG GAC TAT ACG GGC ATT TAC AAA     1014
Leu Val Leu Thr Asn Ala Leu Ile Val Asp Tyr Thr Gly Ile Tyr Lys
        70                  75                  80

GCC GAC ATT GGG ATT AAA GAC GGC AAG ATT GCA GGC ATT GGC AAG GCA     1062
Ala Asp Ile Gly Ile Lys Asp Gly Lys Ile Ala Gly Ile Gly Lys Ala
    85                  90                  95

GGC AAT AAG GAC ATG CAA GAT GGC GTA GAT AAT AAT CTT TGC GTA GGT     1110
Gly Asn Lys Asp Met Gln Asp Gly Val Asp Asn Asn Leu Cys Val Gly
100                 105                 110                 115
```

```
CCT GCT ACA GAG GCT TTG GCA GCT GAG GGC TTG ATT GTA ACC GCT GGT      1158
Pro Ala Thr Glu Ala Leu Ala Ala Glu Gly Leu Ile Val Thr Ala Gly
                120                 125                 130

GGC ATC GAT ACG CAT ATT CAC TTT ATC TCT CCC CAA CAA ATC CCT ACT      1206
Gly Ile Asp Thr His Ile His Phe Ile Ser Pro Gln Gln Ile Pro Thr
            135                 140                 145

GCT TTT GCC AGC GGG GTT ACA ACC ATG ATT GGA GGA GGC ACA GGA CCT      1254
Ala Phe Ala Ser Gly Val Thr Thr Met Ile Gly Gly Gly Thr Gly Pro
        150                 155                 160

GCG GAT GGC ACG AAT GCG ACC ACC ATC ACT CCC GGA CGC GCT AAT CTA      1302
Ala Asp Gly Thr Asn Ala Thr Thr Ile Thr Pro Gly Arg Ala Asn Leu
    165                 170                 175

AAA AGT ATG TTG CGT GCA GCC GAA GAA TAC GCC ATG AAT CTA GGC TTT      1350
Lys Ser Met Leu Arg Ala Ala Glu Glu Tyr Ala Met Asn Leu Gly Phe
180                 185                 190                 195

TTG GCT AAG GGG AAT GTG TCT TAC GAA CCC TCT TTA CGC GAT CAG ATT      1398
Leu Ala Lys Gly Asn Val Ser Tyr Glu Pro Ser Leu Arg Asp Gln Ile
                200                 205                 210

GAA GCA GGG GCG ATT GGT TTT AAA ATC CAC GAA GAC TGG GGA AGC ACA      1446
Glu Ala Gly Ala Ile Gly Phe Lys Ile His Glu Asp Trp Gly Ser Thr
            215                 220                 225

CCT GCA GCT ATT CAC CAC TGC CTC AAT GTC GCC GAT GAA TAC GAT GTG      1494
Pro Ala Ala Ile His His Cys Leu Asn Val Ala Asp Glu Tyr Asp Val
        230                 235                 240

CAA GTG GCT ATC CAC ACC GAT ACC CTT AAC GAG GCG GGC TGT GTA GAA      1542
Gln Val Ala Ile His Thr Asp Thr Leu Asn Glu Ala Gly Cys Val Glu
    245                 250                 255

GAC ACC CTA GAG GCG ATT GCC GGG CGC ACC ATC CAT ACC TTC CAC ACT      1590
Asp Thr Leu Glu Ala Ile Ala Gly Arg Thr Ile His Thr Phe His Thr
260                 265                 270                 275

GAA GGG GCT GGG GGT GGA CAC GCT CCA GAT GTT ATC AAA ATG GCA GGG      1638
Glu Gly Ala Gly Gly Gly His Ala Pro Asp Val Ile Lys Met Ala Gly
                280                 285                 290

GAA TTT AAC ATT CTA CCC GCC TCT ACT AAC CCG ACC ATT CCT TTC ACC      1686
Glu Phe Asn Ile Leu Pro Ala Ser Thr Asn Pro Thr Ile Pro Phe Thr
            295                 300                 305

AAA AAC ACT GAA GCC GAG CAC ATG GAC ATG TTA ATG GTG TGC CAC CAC      1734
Lys Asn Thr Glu Ala Glu His Met Asp Met Leu Met Val Cys His His
        310                 315                 320

TTG GAT AAA AGT ATC AAG GAA GAT GTG CAG TTT GCC GAT TCG AGG ATT      1782
Leu Asp Lys Ser Ile Lys Glu Asp Val Gln Ph  Ala Asp Ser Arg Ile
    325                 330                 335
```

```
CGC CCC CAA ACT ATC GCG GCT GAA GAC CAA CTC CAT GAC ATG GGG ATC     1830
Arg Pro Gln Thr Ile Ala Ala Glu Asp Gln Leu His Asp Met Gly Il
340                 345                 350                 355

TTT TCT ATC ACC AGC TCC GAC TCT CAG GCT ATG GGA CGC GTA GGC GAG     1878
Phe Ser Ile Thr Ser Ser Asp Ser Gln Ala Met Gly Arg Val Gly Glu
                360                 365                 370

GTG ATC ACA CGC ACT TGG CAG ACA GCA GAC AAA AAC AAA AAA GAG TTT     1926
Val Ile Thr Arg Thr Trp Gln Thr Ala Asp Lys Asn Lys Lys Glu Phe
            375                 380                 385

GGG CGC TTG AAA GAG GAA AAA GGC GAT AAC GAC AAC TTC CGC ATC AAA     1974
Gly Arg Leu Lys Glu Glu Lys Gly Asp Asn Asp Asn Phe Arg Ile Lys
        390                 395                 400

CGC TAC ATC TCT AAA TAC ACC ATC AAC CCC GGG ATC GCG CAT GGG ATT     2022
Arg Tyr Ile Ser Lys Tyr Thr Ile Asn Pro Gly Ile Ala His Gly Ile
    405                 410                 415

TCT GAC TAT GTG GGC TCT GTG GAA GTG GGC AAA TAC GCC GAC CTC GTG     2070
Ser Asp Tyr Val Gly Ser Val Glu Val Gly Lys Tyr Ala Asp Leu Val
420                 425                 430                 435

CTT TGG AGT CCG GCT TTC TTT GGC ATT AAG CCC AAT ATG ATT ATT AAG     2118
Leu Trp Ser Pro Ala Phe Phe Gly Ile Lys Pro Asn Met Ile Ile Lys
                440                 445                 450

GGC GGA TTT ATT GCG CTC TCT CAA ATG GGC GAT GCC AAT GCG TCT ATT     2166
Gly Gly Phe Ile Ala Leu Ser Gln Met Gly Asp Ala Asn Ala Ser Ile
            455                 460                 465

CCC ACC CCT CAG CCC GTC TAT TAC CGT GAA ATG TTT GGA CAC CAT GGG     2214
Pro Thr Pro Gln Pro Val Tyr Tyr Arg Glu Met Phe Gly His His Gly
        470                 475                 480

AAA AAC AAA TTC GAC ACC AAT ATC ACT TTC GTG TCC CAA GCG GCT TAC     2262
Lys Asn Lys Phe Asp Thr Asn Ile Thr Phe Val Ser Gln Ala Ala Tyr
    485                 490                 495

AAG GCA GGG ATC AAA GAA GAA CTA GGG CTA GAT CGC GCG GCA CCG CCA     2310
Lys Ala Gly Ile Lys Glu Glu Leu Gly Leu Asp Arg Ala Ala Pro Pro
500                 505                 510                 515

GTG AAA AAC TGT CGC AAT ATC ACT AAA AAG GAC CTC AAA TTC AAC GAT     2358
Val Lys Asn Cys Arg Asn Ile Thr Lys Lys Asp Leu Lys Phe Asn Asp
                520                 525                 530

GTG ACC GCA CAT ATT GAT GTC AAC CCT GAA ACC TAT AAG GTG AAA GTG     2406
Val Thr Ala His Ile Asp Val Asn Pro Glu Thr Tyr Lys Val Lys Val
            535                 540                 545

GAT GGC AAA GAG GTA ACC TCT AAA GCA GCA GAT GAA TTG AGC CTA GCG     2454
Asp Gly Lys Glu Val Thr Ser Lys Ala Ala Asp Glu Leu Ser Leu Ala
        550                 555                 560
```

```
CAA CTT TAT AAT TTG TTC TAGGAGGCTA AGGAGGGGGA TAGAGGGGGT          2502
Gln Leu Tyr Asn Leu Phe
    565                 570

TTATTTAGAG GGGAGTCATT GATTTACCTT TGCTAGTTTA TAATGGATTT AAGAGAGGTT   2562

TTTTTTCGTG TTTTATACCG CGTTGAAACC CTCAAATCTT TACCAAAAGG ATGGTAA      2619
```

(2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 237 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM: Helicobacter felis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

```
Met Lys Leu Thr Pro Lys Glu Leu Asp Lys Leu Met Leu His Tyr Ala
 1               5                  10                  15

Gly Arg Leu Ala Glu Glu Arg Leu Ala Arg Gly Val Lys Leu Asn Tyr
            20                  25                  30

Thr Glu Ala Val Ala Leu Ile Ser Gly Arg Val Met Glu Lys Ala Arg
        35                  40                  45

Asp Gly Asn Lys Ser Val Ala Asp Leu Met Gln Glu Gly Arg Thr Trp
    50                  55                  60

Leu Lys Lys Glu Asn Val Met Asp Gly Val Ala Ser Met Ile His Glu
65                  70                  75                  80

Val Gly Ile Glu Ala Asn Phe Pro Asp Gly Thr Lys Leu Val Thr Ile
                85                  90                  95

His Thr Pro Val Glu Asp Asn Gly Lys Leu Ala Pro Gly Glu Val Phe
            100                 105                 110

Leu Lys Asn Glu Asp Ile Thr Ile Asn Ala Gly Lys Glu Ala Ile Ser
        115                 120                 125

Leu Lys Val Lys Asn Lys Gly Asp Arg Pro Val Gln Val Gly Ser His
    130                 135                 140

Phe His Phe Phe Glu Val Asn Lys Leu Leu Asp Phe Asp Arg Ala Lys
145                 150                 155                 160

Ser Phe Cys Lys Arg Leu Asp Ile Ala Ser Gly Thr Ala Val Arg Phe
                165                 170                 175
```

```
Glu Pro Gly Glu Glu Lys Ser Val Glu Leu Ile Asp Ile Gly Gly Asn
            180                 185                 190

Lys Arg Ile Tyr Gly Phe Asn Ser Leu Val Asp Arg Gln Ala Asp Ala
        195                 200                 205

Asp Gly Lys Lys Leu Gly Leu Lys Arg Ala Lys Glu Lys Gly Phe Gly
    210                 215                 220

Ser Val Asn Cys Gly Cys Glu Ala Thr Lys Asp Lys Gln
225                 230                 235
```

(2) INFORMATION FOR SEQ ID NO: 3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 569 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE :
(A) ORGANISM: Helicobacter felis
(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:

```
Met Lys Lys Ile Ser Arg Lys Glu Tyr Val Ser Met Tyr Gly Pro Thr
  1               5                  10                  15

Thr Gly Asp Arg Val Arg Leu Gly Asp Thr Asp Leu Ile Leu Glu Val
             20                  25                  30

Glu His Asp Cys Thr Thr Tyr Gly Glu Glu Ile Lys Phe Gly Gly Gly
         35                  40                  45

Lys Thr Ile Arg Asp Gly Met Ser Gln Thr Asn Ser Pro Ser Ser Tyr
     50                  55                  60

Glu Leu Asp Leu Val Leu Thr Asn Ala Leu Ile Val Asp Tyr Thr Gly
 65                  70                  75                  80

Ile Tyr Lys Ala Asp Ile Gly Ile Lys Asp Gly Lys Ile Ala Gly Ile
                 85                  90                  95

Gly Lys Ala Gly Asn Lys Asp Met Gln Asp Gly Val Asp Asn Asn Leu
            100                 105                 110

Cys Val Gly Pro Ala Thr Glu Ala Leu Ala Ala Glu Gly Leu Ile Val
        115                 120                 125

Thr Ala Gly Gly Ile Asp Thr His Ile His Phe Ile Ser Pro Gln Gln
    130                 135                 140

Ile Pro Thr Ala Phe Ala Ser Gly Val Thr Thr Met Ile Gly Gly Gly
145                 150                 155                 160
```

```
Thr Gly Pro Ala Asp Gly Thr Asn Ala Thr Thr Ile Thr Pro Gly Arg
                165                 170                 175

Ala Asn Leu Lys Ser Met Leu Arg Ala Ala Glu Glu Tyr Ala Met Asn
            180                 185                 190

Leu Gly Phe Leu Ala Lys Gly Asn Val Ser Tyr Glu Pro Ser Leu Arg
        195                 200                 205

Asp Gln Ile Glu Ala Gly Ala Ile Gly Phe Lys Ile His Glu Asp Trp
    210                 215                 220

Gly Ser Thr Pro Ala Ala Ile His His Cys Leu Asn Val Ala Asp Glu
225                 230                 235                 240

Tyr Asp Val Gln Val Ala Ile His Thr Asp Thr Leu Asn Glu Ala Gly
                245                 250                 255

Cys Val Glu Asp Thr Leu Glu Ala Ile Ala Gly Arg Thr Ile His Thr
            260                 265                 270

Phe His Thr Glu Gly Ala Gly Gly Gly His Ala Pro Asp Val Ile Lys
        275                 280                 285

Met Ala Gly Glu Phe Asn Ile Leu Pro Ala Ser Thr Asn Pro Thr Ile
    290                 295                 300

Pro Phe Thr Lys Asn Thr Glu Ala Glu His Met Asp Met Leu Met Val
305                 310                 315                 320

Cys His His Leu Asp Lys Ser Ile Lys Glu Asp Val Gln Phe Ala Asp
                325                 330                 335

Ser Arg Ile Arg Pro Gln Thr Ile Ala Ala Glu Asp Gln Leu His Asp
            340                 345                 350

Met Gly Ile Phe Ser Ile Thr Ser Ser Asp Ser Gln Ala Met Gly Arg
        355                 360                 365

Val Gly Glu Val Ile Thr Arg Thr Trp Gln Thr Ala Asp Lys Asn Lys
    370                 375                 380

Lys Glu Phe Gly Arg Leu Lys Glu Glu Lys Gly Asp Asn Asp Asn Phe
385                 390                 395                 400

Arg Ile Lys Arg Tyr Ile Ser Lys Tyr Thr Ile Asn Pro Gly Ile Ala
                405                 410                 415

His Gly Ile Ser Asp Tyr Val Gly Ser Val Glu Val Gly Lys Tyr Ala
            420                 425                 430

Asp Leu Val Leu Trp Ser Pro Ala Phe Phe Gly Ile Lys Pro Asn Met
        435                 440                 445

Ile Ile Lys Gly Gly Phe Ile Ala Leu Ser Gln Met Gly Asp Ala Asn
    450                 455                 460
```

```
Ala Ser Ile Pro Thr Pro Gln Pro Val Tyr Tyr Arg Glu Met Phe Gly
465                 470                 475                 480

His His Gly Lys Asn Lys Phe Asp Thr Asn Ile Thr Phe Val Ser Gln
                485                 490                 495

Ala Ala Tyr Lys Ala Gly Ile Lys Glu Glu Leu Gly Leu Asp Arg Ala
            500                 505                 510

Ala Pro Pro Val Lys Asn Cys Arg Asn Ile Thr Lys Lys Asp Leu Lys
        515                 520                 525

Phe Asn Asp Val Thr Ala His Ile Asp Val Asn Pro Glu Thr Tyr Lys
    530                 535                 540

Val Lys Val Asp Gly Lys Glu Val Thr Ser Lys Ala Ala Asp Glu Leu
545                 550                 555                 560

Ser Leu Ala Gln Leu Tyr Asn Leu Phe
                565
```

(2) INFORMATION FOR SEQ ID NO: 4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 2284 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 124..477
(D) OTHER INFORMATION: /standard_name= "H. pylori - Hsp A"

(ix) FEATURE:
(A) NAME/KEY: CDS
(B) LOCATION: 506..2143
(D) OTHER INFORMATION: /standard_name= "H. pylori - Hsp B"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

```
ACAAACATGA TCTCATATCA GGGACTTGTT CGCACCTTCC CTAAAAATGC GCTATAGTTG     60

TGTCGCTTAA GAATACTAAG CGCTAAATTT CTATTTTATT TATCAAAACT TAGGAGAACT    120

GAA ATG AAG TTT CAA CCA TTA GGA GAA AGG GTC TTA GTA GAA AGA CTT      168
    Met Lys Phe Gln Pro Leu Gly Glu Arg Val Leu Val Glu Arg Leu
      1               5                  10                  15

GAA GAA GAG AAC AAA ACC AGT TCA GGC ATC ATC ATC CCT GAT AAC GCT      216
Glu Glu Glu Asn Lys Thr Ser Ser Gly Ile Ile Ile Pr  Asp Asn Ala
                20                  25                  30
```

```
AAA GAA AAG CCT TTA ATG GGC GTA GTC AAA GCG GTT AGC CAT AAA ATC      264
Lys Glu Lys Pro Leu Met Gly Val Val Lys Ala Val Ser His Lys Ile
            35                  40                  45

AGT GAG GGT TGC AAA TGC GTT AAA GAA GGC GAT GTG ATC GCT TTT GGC      312
Ser Glu Gly Cys Lys Cys Val Lys Glu Gly Asp Val Ile Ala Phe Gly
        50                  55                  60

AAA TAC AAA GGC GCA GAA ATC GTT TTA GAT GGC GTT GAA TAC ATG GTG      360
Lys Tyr Lys Gly Ala Glu Ile Val Leu Asp Gly Val Glu Tyr Met Val
    65                  70                  75

CTA GAA CTA GAA GAC ATT CTA GGT ATT GTG GGC TCA GGC TCT TGC TGT      408
Leu Glu Leu Glu Asp Ile Leu Gly Ile Val Gly Ser Gly Ser Cys Cys
80                  85                  90                  95

CAT ACA GGT AAT CAT GAT CAT AAA CAT GCT AAA GAG CAT GAA GCT TGC      456
His Thr Gly Asn His Asp His Lys His Ala Lys Glu His Glu Ala Cys
                100                 105                 110

TGT CAT GAT CAC AAA AAA CAC TAAAAAACAT TATTATTAAG GATACAAA ATG       508
Cys His Asp His Lys Lys His                                Met
            115                                              1

GCA AAA GAA ATC AAA TTT TCA GAT AGC GCA AGA AAC CTT TTA TTT GAA      556
Ala Lys Glu Ile Lys Phe Ser Asp Ser Ala Arg Asn Leu Leu Phe Glu
            5                   10                  15

GGC GTA AGA CAA CTC CAT GAC GCT GTC AAA GTA ACC ATG GGC CCA AGA      604
Gly Val Arg Gln Leu His Asp Ala Val Lys Val Thr Met Gly Pro Arg
        20                  25                  30

GGC AGG AAC GTG TTG ATC CAA AAA AGC TAT GGC GCT CCA AGC ATC ACC      652
Gly Arg Asn Val Leu Ile Gln Lys Ser Tyr Gly Ala Pro Ser Ile Thr
    35                  40                  45

AAA GAC GGC GTG AGC GTG GCT AAA GAG ATT GAA TTA AGT TGC CCC GTG      700
Lys Asp Gly Val Ser Val Ala Lys Glu Ile Glu Leu Ser Cys Pro Val
50                  55                  60                  65

GCT AAC ATG GGC GCT CAG CTC GTT AAA GAA GAT GCG AGC AAA ACC GCT      748
Ala Asn Met Gly Ala Gln Leu Val Lys Glu Asp Ala Ser Lys Thr Ala
                70                  75                  80

GAT GCC GCC GGC GAT GGC ACG ACC ACA GCG ACC GTG CTG GCT TAT AGC      796
Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Tyr Ser
            85                  90                  95

ATT TTT AAA GAG GGC TTG AGG AAT ATC ACG GCT GGG GCT AAC CCT ATT      844
Ile Phe Lys Glu Gly Leu Arg Asn Ile Thr Ala Gly Ala Asn Pro Ile
        100                 105                 110

GAA GTG AAA CGA GGC ATG GAT AAA GCG CCT GAA GCG ATC ATT AAT GAG      892
Glu Val Lys Arg Gly Met Asp Lys Ala Pro Glu Ala Ile Ile Asn Glu
    115                 120                 125
```

```
CTT AAA AAA GCG AGC AAA AAA GTG GGC GGT AAA GAA GAA ATC ACC CAA      940
Leu Lys Lys Ala Ser Lys Lys Val Gly Gly Lys Glu Glu Ile Thr Gln
130                 135                 140                 145

GTA GCG ACC ATT TCT GCA AAC TCC GAT CAC AAT ATC GGG AAA CTC ATC      988
Val Ala Thr Ile Ser Ala Asn Ser Asp His Asn Ile Gly Lys Leu Ile
                150                 155                 160

GCT GAC GCT ATG GAA AAA GTG GGT AAA GAC GGC GTG ATC ACC GTT GAA     1036
Ala Asp Ala Met Glu Lys Val Gly Lys Asp Gly Val Ile Thr Val Glu
            165                 170                 175

GAA GCT AAG GGC ATT GAA GAT GAA TTA GAT GTC GTA GAA GGC ATG CAA     1084
Glu Ala Lys Gly Ile Glu Asp Glu Leu Asp Val Val Glu Gly Met Gln
        180                 185                 190

TTT GAT AGA GGC TAC CTC TCC CCT TAC TTT GTA ACC AAC GCT GAG AAA     1132
Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Val Thr Asn Ala Glu Lys
    195                 200                 205

ATG ACC GCT CAA TTG GAT AAC GCT TAC ATC CTT TTA ACG GAT AAA AAA     1180
Met Thr Ala Gln Leu Asp Asn Ala Tyr Ile Leu Leu Thr Asp Lys Lys
210                 215                 220                 225

ATC TCT AGC ATG AAA GAC ATT CTC CCG CTA CTA GAA AAA ACC ATG AAA     1228
Ile Ser Ser Met Lys Asp Ile Leu Pro Leu Leu Glu Lys Thr Met Lys
                230                 235                 240

GAG GGC AAA CCG CTT TTA ATC ATC GCT GAA GAC ATT GAG GGC GAA GCT     1276
Glu Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Ile Glu Gly Glu Ala
            245                 250                 255

TTA ACG ACT CTA GTG GTG AAT AAA TTA AGA GGC GTG TTG AAT ATC GCA     1324
Leu Thr Thr Leu Val Val Asn Lys Leu Arg Gly Val Leu Asn Ile Ala
        260                 265                 270

GCG GTT AAA GCT CCA GGC TTT GGG GAC AGG AGA AAA GAA ATG CTC AAA     1372
Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Glu Met Leu Lys
    275                 280                 285

GAC ATC GCT GTT TTA ACC GGC GGT CAA GTC ATT AGC GAA GAA TTG GGC     1420
Asp Ile Ala Val Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Leu Gly
290                 295                 300                 305

TTG AGT CTA GAA AAC GCT GAA GTG GAG TTT TTA GGC AAA GCG AAG ATT     1468
Leu Ser Leu Glu Asn Ala Glu Val Glu Phe Leu Gly Lys Ala Lys Ile
                310                 315                 320

GTG ATT GAC AAA GAC AAC ACC ACG ATC GTA GAT GGC AAA GGC CAT AGC     1516
Val Ile Asp Lys Asp Asn Thr Thr Ile Val Asp Gly Lys Gly His Ser
            325                 330                 335

CAT GAC GTC AAA GAC AGA GTC GCG CAA ATC AAA ACC CAA ATT GCA AGC     1564
His Asp Val Lys Asp Arg Val Ala Gln Ile Lys Thr Gln Ile Ala Ser
        340                 345                 350
```

```
ACG ACA AGC GAT TAC GAC AAA GAA AAA TTG CAA GAA AGA TTG GCC AAA      1612
Thr Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala Lys
    355                 360                 365

CTC TCT GGC GGT GTG GCT GTG ATT AAA GTG GGC GCT GCG AGT GAA GTG      1660
Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Ser Glu Val
370                 375                 380                 385

GAA ATG AAA GAG AAA AAA GAC CGG GTG GAT GAC GCG TTG AGC GCG ACT      1708
Glu Met Lys Glu Lys Lys Asp Arg Val Asp Asp Ala Leu Ser Ala Thr
                390                 395                 400

AAA GCG GCG GTT GAA GAA GGC ATT GTG ATT GGG GGC GGT GCG GCC CTC      1756
Lys Ala Ala Val Glu Glu Gly Ile Val Ile Gly Gly Gly Ala Ala Leu
            405                 410                 415

ATT CGC GCG GCC CAA AAA GTG CAT TTG AAT TTA CAC GAT GAT GAA AAA      1804
Ile Arg Ala Ala Gln Lys Val His Leu Asn Leu His Asp Asp Glu Lys
        420                 425                 430

GTG GGC TAT GAA ATC ATC ATG CGC GCC ATT AAA GCC CCA TTA GCT CAA      1852
Val Gly Tyr Glu Ile Ile Met Arg Ala Ile Lys Ala Pro Leu Ala Gln
    435                 440                 445

ATC GCT ATC AAT GCC GGT TAT GAT GGC GGT GTG GTC GTG AAT GAA GTA      1900
Ile Ala Ile Asn Ala Gly Tyr Asp Gly Gly Val Val Val Asn Glu Val
450                 455                 460                 465

GAA AAA CAC GAA GGG CAT TTT GGT TTT AAC GCT AGC AAT GGC AAG TAT      1948
Glu Lys His Glu Gly His Phe Gly Phe Asn Ala Ser Asn Gly Lys Tyr
                470                 475                 480

GTG GAC ATG TTT AAA GAA GGC ATT ATT GAC CCC TTA AAA GTA GAA AGG      1996
Val Asp Met Phe Lys Glu Gly Ile Ile Asp Pro Leu Lys Val Glu Arg
            485                 490                 495

ATC GCT TTA CAA AAT GCG GTT TCG GTT TCA AGC CTG CTT TTA ACC ACA      2044
Ile Ala Leu Gln Asn Ala Val Ser Val Ser Ser Leu Leu Leu Thr Thr
        500                 505                 510

GAA GCC ACC GTG CAT GAA ATC AAA GAA GAA AAA GCG GCC CCA GCA ATG      2092
Glu Ala Thr Val His Glu Ile Lys Glu Glu Lys Ala Ala Pro Ala Met
    515                 520                 525

CCT GAT ATG GGT GGC ATG GGC GGA ATG GGA GGC ATG GGC GGC ATG ATG      2140
Pro Asp Met Gly Gly Met Gly Gly Met Gly Gly Met Gly Gly Met Met
530                 535                 540                 545

TAAGCCCCCT TGCTTTTTGG TATCATCTGC TTTTAAAATC CATCTTCTAG AATCCCCCCT    2200

TCTAAAATCC CTTTTTTGGG GGGTGCTTTT GGTTTGATAA AACCGCTCGC TTTTAAAAAC    2260

GCGCAACAAA AAACTCTGTT AAGC                                           2284
```

(2) INFORMATION FOR SEQ ID NO: 5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 545 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM : H. pylori

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

```
Met Ala Lys Glu Ile Lys Phe Ser Asp Ser Ala Arg Asn Leu Leu Phe
 1               5                  10                  15

Glu Gly Val Arg Gln Leu His Asp Ala Val Lys Val Thr Met Gly Pro
            20                  25                  30

Arg Gly Arg Asn Val Leu Ile Gln Lys Ser Tyr Gly Ala Pro Ser Ile
        35                  40                  45

Thr Lys Asp Gly Val Ser Val Ala Lys Glu Ile Glu Leu Ser Cys Pro
    50                  55                  60

Val Ala Asn Met Gly Ala Gln Leu Val Lys Glu Asp Ala Ser Lys Thr
65                  70                  75                  80

Ala Asp Ala Ala Gly Asp Gly Thr Thr Thr Ala Thr Val Leu Ala Tyr
                85                  90                  95

Ser Ile Phe Lys Glu Gly Leu Arg Asn Ile Thr Ala Gly Ala Asn Pro
            100                 105                 110

Ile Glu Val Lys Arg Gly Met Asp Lys Ala Pro Glu Ala Ile Ile Asn
        115                 120                 125

Glu Leu Lys Lys Ala Ser Lys Lys Val Gly Gly Lys Glu Glu Ile Thr
    130                 135                 140

Gln Val Ala Thr Ile Ser Ala Asn Ser Asp His Asn Ile Gly Lys Leu
145                 150                 155                 160

Ile Ala Asp Ala Met Glu Lys Val Gly Lys Asp Gly Val Ile Thr Val
                165                 170                 175

Glu Glu Ala Lys Gly Ile Glu Asp Glu Leu Asp Val Val Glu Gly Met
            180                 185                 190

Gln Phe Asp Arg Gly Tyr Leu Ser Pro Tyr Phe Val Thr Asn Ala Glu
        195                 200                 205

Lys Met Thr Ala Gln Leu Asp Asn Ala Tyr Ile Leu Leu Thr Asp Lys
    210                 215                 220
```

```
Lys Ile Ser Ser Met Lys Asp Ile Leu Pro Leu Leu Glu Lys Thr Met
225                 230                 235                 240

Lys Glu Gly Lys Pro Leu Leu Ile Ile Ala Glu Asp Ile Glu Gly Glu
                245                 250                 255

Ala Leu Thr Thr Leu Val Val Asn Lys Leu Arg Gly Val Leu Asn Ile
            260                 265                 270

Ala Ala Val Lys Ala Pro Gly Phe Gly Asp Arg Arg Lys Glu Met Leu
        275                 280                 285

Lys Asp Ile Ala Val Leu Thr Gly Gly Gln Val Ile Ser Glu Glu Leu
    290                 295                 300

Gly Leu Ser Leu Glu Asn Ala Glu Val Glu Phe Leu Gly Lys Ala Lys
305                 310                 315                 320

Ile Val Ile Asp Lys Asp Asn Thr Thr Ile Val Asp Gly Lys Gly His
                325                 330                 335

Ser His Asp Val Lys Asp Arg Val Ala Gln Ile Lys Thr Gln Ile Ala
            340                 345                 350

Ser Thr Thr Ser Asp Tyr Asp Lys Glu Lys Leu Gln Glu Arg Leu Ala
        355                 360                 365

Lys Leu Ser Gly Gly Val Ala Val Ile Lys Val Gly Ala Ala Ser Glu
    370                 375                 380

Val Glu Met Lys Glu Lys Lys Asp Arg Val Asp Asp Ala Leu Ser Ala
385                 390                 395                 400

Thr Lys Ala Ala Val Glu Glu Gly Ile Val Ile Gly Gly Gly Ala Ala
                405                 410                 415

Leu Ile Arg Ala Ala Gln Lys Val His Leu Asn Leu His Asp Asp Glu
            420                 425                 430

Lys Val Gly Tyr Glu Ile Ile Met Arg Ala Ile Lys Ala Pro Leu Ala
        435                 440                 445

Gln Ile Ala Ile Asn Ala Gly Tyr Asp Gly Gly Val Val Val Asn Glu
    450                 455                 460

Val Glu Lys His Glu Gly His Phe Gly Phe Asn Ala Ser Asn Gly Lys
465                 470                 475                 480

Tyr Val Asp Met Phe Lys Glu Gly Ile Ile Asp Pro Leu Lys Val Glu
                485                 490                 495

Arg Ile Ala Leu Gln Asn Ala Val Ser Val Ser Ser Leu Leu Leu Thr
            500                 505                 510

Thr Glu Ala Thr Val His Glu Ile Lys Glu Glu Lys Ala Ala Pro Ala
        515                 520                 525
```

```
Met Pro Asp Met Gly Gly Met Gly Gly Met Gly Gly Met Gly Gly Met
    530                 535                 540

Met
545
```

(2) INFORMATION FOR SEQ ID NO: 6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 118 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM : H. pylori

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:

```
Met Lys Phe Gln Pro Leu Gly Glu Arg Val Leu Val Glu Arg Leu Glu
  1               5                  10                  15

Glu Glu Asn Lys Thr Ser Ser Gly Ile Ile Ile Pro Asp Asn Ala Lys
             20                  25                  30

Glu Lys Pro Leu Met Gly Val Val Lys Ala Val Ser His Lys Ile Ser
         35                  40                  45

Glu Gly Cys Lys Cys Val Lys Glu Gly Asp Val Ile Ala Phe Gly Lys
     50                  55                  60

Tyr Lys Gly Ala Glu Ile Val Leu Asp Gly Val Glu Tyr Met Val Leu
 65                  70                  75                  80

Glu Leu Glu Asp Ile Leu Gly Ile Val Gly Ser Gly Ser Cys Cys His
                 85                  90                  95

Thr Gly Asn His Asp His Lys His Ala Lys Glu His Glu Ala Cys Cys
            100                 105                 110

His Asp His Lys Lys His
        115
```

(2) INFORMATION FOR SEQ ID NO: 7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 591 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

```
(vi) ORIGINAL SOURCE:
     (A) ORGANISM : H. felis

(ix) FEATURE:
     (A) NAME/KEY: CDS
     (B) LOCATION: 1..591
     (D) OTHER INFORMATION: /standard_name= "URE I"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:

ATG TTA GGT CTT GTG TTA TTG TAT GTT GCG GTC GTG CTG ATC AGC AAC      48
Met Leu Gly Leu Val Leu Leu Tyr Val Ala Val Val Leu Ile Ser Asn
  1               5                  10                  15

GGA GTT AGT GGG CTT GCA AAT GTG GAT GCC AAA AGC AAA GCC ATC ATG      96
Gly Val Ser Gly Leu Ala Asn Val Asp Ala Lys Ser Lys Ala Ile Met
             20                  25                  30

AAC TAC TTT GTG GGG GGG GAC TCT CCA TTG TGT GTA ATG TGG TCG CTA     144
Asn Tyr Phe Val Gly Gly Asp Ser Pro Leu Cys Val Met Trp Ser Leu
         35                  40                  45

TCA TCT TAT TCC ACT TTC CAC CCC ACC CCC CCT GCA ACT GGT CCA GAA     192
Ser Ser Tyr Ser Thr Phe His Pro Thr Pro Pro Ala Thr Gly Pro Glu
     50                  55                  60

GAT GTC GCG CAG GTG TCT CAA CAC CTC ATT AAC TTC TAT GGT CCA GCG     240
Asp Val Ala Gln Val Ser Gln His Leu Ile Asn Phe Tyr Gly Pro Ala
 65                  70                  75                  80

ACT GGT CTA TTG TTT GGT TTT ACC TAC TTG TAT GCT GCC ATC AAC AAC     288
Thr Gly Leu Leu Phe Gly Phe Thr Tyr Leu Tyr Ala Ala Ile Asn Asn
                 85                  90                  95

ACT TTC AAT CTC GAT TGG AAA CCC TAT GGC TGG TAT TGC TTG TTT GTA     336
Thr Phe Asn Leu Asp Trp Lys Pro Tyr Gly Trp Tyr Cys Leu Phe Val
            100                 105                 110

ACC ATC AAC ACT ATC CCA GCG GCC ATT CTT TCT CAC TAT TCC GAT GCG     384
Thr Ile Asn Thr Ile Pro Ala Ala Ile Leu Ser His Tyr Ser Asp Ala
        115                 120                 125

CTT GAT GAT CAC CGC CTC TTA GGA ATC ACT GAG GGC GAT TGG TGG GCT     432
Leu Asp Asp His Arg Leu Leu Gly Ile Thr Glu Gly Asp Trp Trp Ala
    130                 135                 140

TTC ATT TGG CTT GCT TGG GGT GTT TTG TGG CTC ACT GGT TGG ATT GAA     480
Phe Ile Trp Leu Ala Trp Gly Val Leu Trp Leu Thr Gly Trp Ile Glu
145                 150                 155                 160

TGC GCA CTT GGT AAG AGT CTA GGT AAA TTT GTT CCA TGG CTT GCC ATC     528
Cys Ala Leu Gly Lys Ser Leu Gly Lys Phe Val Pro Trp Leu Ala Ile
                165                 170                 175
```

```
GTC GAG GGC GTG ATC ACC GCT TGG ATT CCT GCT TGG CTA CTC TTT ATC       576
Val Glu Gly Val Ile Thr Ala Trp Ile Pro Ala Trp Leu Leu Phe Ile
            180                 185                 190

CAA CAC TGG TCT TGA
591
Gln His Trp Ser
    195
```

(2) INFORMATION FOR SEQ ID NO: 8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 199 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(vi) ORIGINAL SOURCE:
(A) ORGANISM : H. felis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:

```
Lys Gly Trp Met Leu Gly Leu Val Leu Leu Tyr Val Ala Val Val Leu
1               5                   10                  15

Ile Ser Asn Gly Val Ser Gly Leu Ala Asn Val Asp Ala Lys Ser Lys
            20                  25                  30

Ala Ile Met Asn Tyr Phe Val Gly Gly Asp Ser Pro Leu Cys Val Met
        35                  40                  45

Trp Ser Leu Ser Ser Tyr Ser Thr Phe His Pro Thr Pro Pro Ala Thr
    50                  55                  60

Gly Pro Glu Asp Val Ala Gln Val Ser Gln His Leu Ile Asn Phe Tyr
65                  70                  75                  80

Gly Pro Ala Thr Gly Leu Leu Phe Gly Phe Thr Tyr Leu Tyr Ala Ala
                85                  90                  95

Ile Asn Asn Thr Phe Asn Leu Asp Trp Lys Pro Tyr Gly Trp Tyr Cys
            100                 105                 110

Leu Phe Val Thr Ile Asn Thr Ile Pro Ala Ala Ile Leu Ser His Tyr
        115                 120                 125

Ser Asp Ala Leu Asp Asp His Arg Leu Leu Gly Ile Thr Glu Gly Asp
    130                 135                 140

Trp Trp Ala Ph  Ile Trp Leu Ala Trp Gly Val Leu Trp Leu Thr Gly
145                 150                 155                 160

Trp Ile Glu Cys Ala Leu Gly Lys Ser Leu Gly Lys Ph  Val Pro Trp
                165                 170                 175
```

```
Leu Ala Ile Val Glu Gly Val Ile Thr Ala Trp Il  Pro Ala Trp Leu
            180                     185                     190

Leu Phe Ile Gln His Trp Ser
        195
```

CLAIMS

1. Immunogenic composition, capable of inducing antibodies against Helicobacter infection, characterised in that it comprises :

i) at least one sub-unit of a urease structural polypeptide from Helicobacter pylori, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter felis urease, and/or at least one sub-unit of a urease structural polypeptide from Helicobacter felis, or a fragment thereof, said fragment being recognised by antibodies reacting with Helicobacter pylori urease ;

ii) and/or, a Heat Shock protein (HSP), or chaperonin, from Helicobacter , or a fragment of said protein.

2. Immunogenic composition according to claim 1 capable of inducing protective antibodies.

3. Immunogenic composition according to claim 1 characterised in that it includes component (i), which comprises or consists of the Helicobacter felis urease structural polypeptide(s) encoded by the ure A and/or ure B genes of plasmid pILL205 (CNCM I-1355), a polypeptide exhibiting at least 90 % homology with the said polypeptide(s), or a fragment thereof having at least 6 amino-acids and being recognised by antibodies reacting with Helicobacter pylori urease.

4. Immunogenic composition according to claim 1, characterised in that it includes component ii) which is a HSP from Helicobacter pylori, or a fragment thereof.

5. Immunogenic composition according to any of preceding claims characterised in that the HSP is HSP A and/or HSP B encoded by the hsp A and/or hsp B genes respectively, of plasmid pILL689 (CNCM I-1356), or a

polypeptide exhibiting at least 75 % homology with the said HSP's, or a fragment of either or both of these proteins having at least 6 amino-acids.

6. Pharmaceutical composition for use as a vaccine in protecting against Helicobacter infection, particularly against Helicobacter pylori and Helicobacter felis, characterised in that it comprises the immunogenic composition of any of claims 1-5, in combination with physiologically acceptable excipient(s) and possibly adjuvants.

7. Proteinaceous material characterised in that it comprises at least one of the Helicobacter felis polypeptides encoded by the urease gene cluster of the plasmid pILL205 (CNCM I-1355), including the structural and accessory urease polypeptides, or a polypeptide having at least 90 % homology with said polypeptides, or a fragment thereof.

8. Proteinaceous material according to claim 7, characterised in that it consists of or comprises the gene product of ure A and/or ure B as illustrated in figure 3, or a fragment having at least 6 amino-acids, or a variant of these gene products having at least 90 % homology, said fragment and said variant being recognised by antibodies reacting with Helicobacter pylori urease.

9. Proteinaceous material according to claim 7 characterised in that it consists of or comprises the gene product of ure I, as illustrated in figure 9, or a fragment thereof having at least 6 amino-acids, or a variant of the gene product having at least 75 % homology, said fragment and said variant having the capacity to activate the ure A and ure B gene products in the presence of the remaining urease "accessory" gene products.

10. Nucleic acid sequence characterised in that it comprises :

(i) at least one sequence coding for the proteinaceous material of any one of claims 6-9 ;

or (ii) a sequence complementary to sequence (i) ;

or (iii) a sequence capable of hybridising to sequences (i) or (ii) under stringent conditions ;

iv) a fragment of any of sequences (i), (ii) or (iii) comprising at least 10 consecutive nucleotides.

11. Nucleic acid sequence according to claim 9 characterised in that it comprises the sequence of plasmid pILL205 (CNCM I-1355), for example the sequence of Figure 3, in particular that coding for the gene product of ure A and for ure B or the sequence of Figure 9 (Ure I), or a sequence capable of hybridising to these sequences under stringent conditions, or a sequence complementary to these sequences, or a fragment comprising at least 10 consecutive nucleotides of these sequences.

12. Expression vector characterised in that it contains a nucleic acid sequence according to claim 10 or 11.

13. Plasmid pILL205 (CNCM I-1355).

14. Oligonucleotide suitable for use as a primer in a nucleic acid amplification reaction, characterised in that it comprises from 10 to 100 consecutive nucleotides of the sequence of claim 10 or 11.

15. Nucleotide probe characterised in that it comprises a sequence according to any one of claims 9 or 10, with an appropriate labelling means.

16. Prokaryotic or eukaryotic host cell stably transformed by an expression vector according to claim 12 or 13.

17. Proteinaceous material characterised in that it comprises at least one of the Heat Shock Proteins (HSP), or chaperonins, of Helicobacter pylori, or a fragment thereof.

18. Proteinaceous material according to claim 17, characterised in that it comprises or consists of HSP A and/or HSP B, having the amino-acid sequence illustrated in Figure 6, or a polypeptide having at least 75 %, and preferably at least 80 % homology with said polypeptide, or a fragment thereof, comprising at least 6 amino-acids.

19. Proteinaceous material according to claim 18 characterised in that it comprises or consists of the HSP A C-terminal sequence :

G S C C H T G N H D H K H A K E H E A C C H D H K K H

or a fragment comprising at least 6 consecutive amino-acids of this sequence.

20. Nucleic acid sequence characterised in that it comprises :

i) a sequence coding for the proteinaceous material of any one of claims 17 to 19 or of any one of the proteinaceous materials of claims 7 to 9 ;

or ii) a sequence complementary to sequence (i) ;

or iii) a sequence capable of hybridizing to sequence (i) or (ii) under stringent conditions ;

or iv) a fragment of any of sequences (i), (ii) or (iii) comprising at least 10 nucleotides.

21. Nucleic acid sequence according to claim 20 characterised in that it comprises all or part of the sequence of plasmid pILL689 (CNCM I-1356), for example the sequence of figure 6, in particular that coding for HSP A and/or HSP B, or a sequence complementary to this sequence, or a sequence capable of hybridizing to this sequence under stringent conditions, or a fragment thereof.

22. Expression vector characterised in that it contains a nucleic acid sequence according to claim 20 or 21.

23. Plasmid pILL689 (CNCM I-1356).

24. Oligonucleotide suitable for use as a primer in a nucleic acid amplification reaction, characterised in that it comprises from 10 to 100 consecutive nucleotides of the sequence of claim 20 or 21.

25. Nucleotide probe, characterised in that it comprises a sequence according to any one of claims 20 or 21 with an appropriate labelling means.

26. Microorganism, stably transformed by an expression vector according to claim 22 or 23.

27. Monoclonal or polyclonal antibodies or fragments thereof, to the proteinaceous material of any one of claims 8 to 10, characterised in that they are either specific for the Helicobacter felis material, or alternatively, cross-react with the gene products of the urease gene cluster of Helicobacter pylori.

28. Monoclonal or polyclonal antibodies according to claim 27 characterised in that they recognise both the Helicobacter felis ure A and/or ure B gene product, and the Helicobacter pylori ure A and/or ure B gene product.

29. Monoclonal or polyclonal antibodies or fragments thereof, to the proteinaceous material of claims 17 or 18, characterised in that they are either specific for the Helicobacter pylori material or, alternatively, cross-react with GroEL-like proteins or GroES-like proteins from bacteria other than Helicobacter.

30. Monoclonal or polyclonal antibodies according to claim 29 characterised in that they recognise specifically the HSP A C-terminal sequence.

31. Use of the immunogenic composition of claim 1 for the preparation of a vaccine suitable for use in man and animals against Helicobacter infection, particularly against Helicobacter pylori and Helicobacter felis.

32. Use of the antibodies of claims 27 to 30 in a therapeutic composition for treating infection by Helicobacter, in particular Helicobacter pylori, Helicobacter heilmannii and Helicobacter felis in man or animals.

33. Method for the production of a pharmaceutical composition according to claim 6, characterised by culturing a transformed micro-organism according to claim 16, and optionally, also a micro-organism according to claim 26, collecting and purifying the Helicobacter urease polypeptide material and where applicable, also the HSP material, and combining these materials with suitable excipients, adjuvants and, optionally, other additives.

34. Use of nucleotide sequences of any claim 15 or 25 for the in vitro detection in a biological sample, of an infection by Helicobacter, optionally following a gene amplification reaction.

35. Kit for the in vitro detection of Helicobacter infection, characterised in that it comprises :

- a nucleotide probe according to claim 15 or 25 ;

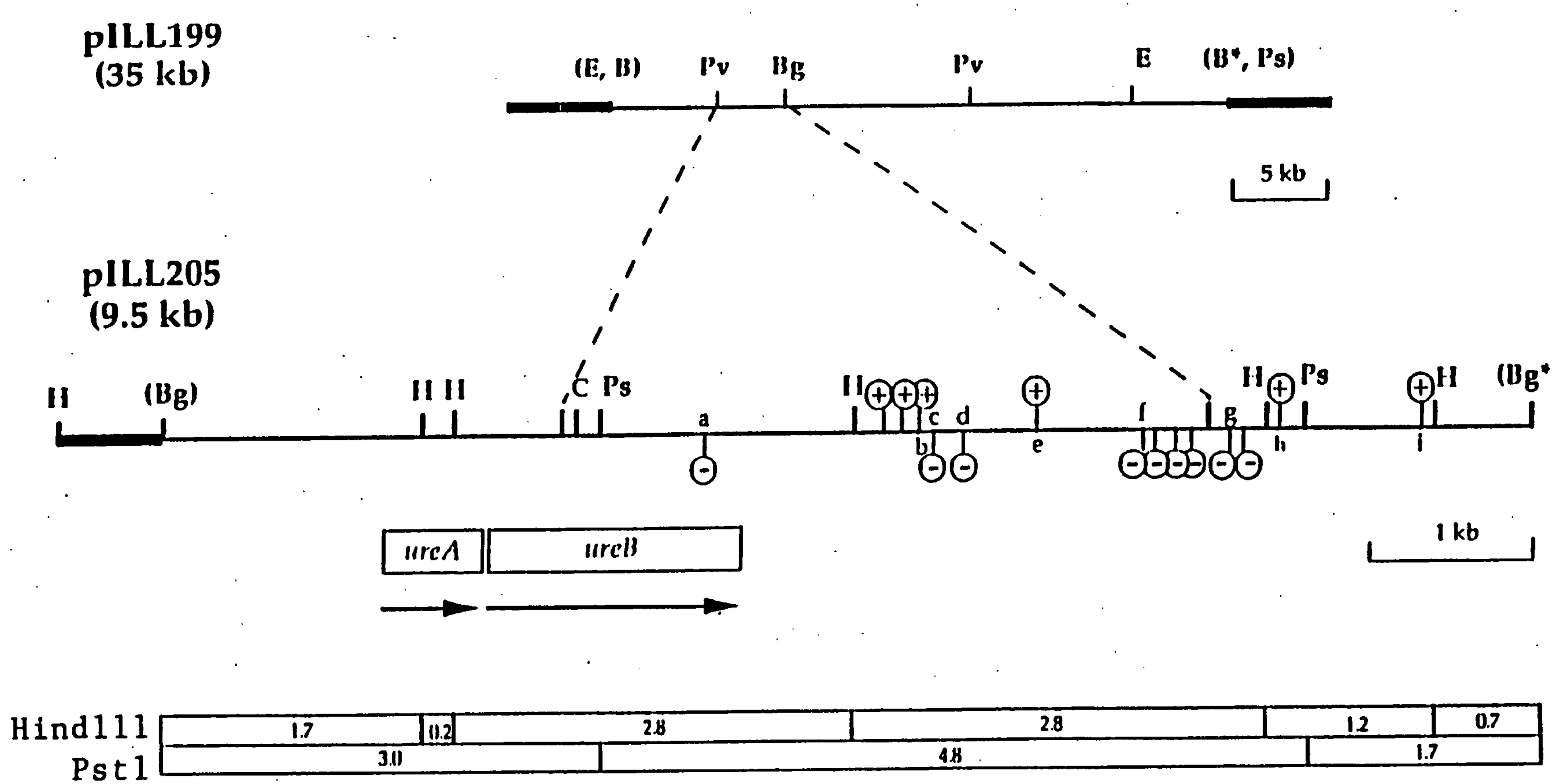
- an appropriate medium for carrying out a hybridisation reaction between the nucleic acid of Helicobacter and the probe ;

- reagents for the detection of any hybrids formed.

36. Proteinaceous material characterised in that it comprises a fusion or mixed protein including at least one sub-unit of a urease structural polypeptide from Helicobacter pylori or fragment thereof, or from Helicobacter felis or fragment thereof as defined in claims 1 to 3, 5, 7 to 9, and or a heat shock proteins (HSP) from Helicobacter or fragment thereof, as defined in claims 17 to 20.

37. Purified antibodies or serum obtained by immunisation of an animal with the immunogenic composition according to claims 1 to 5, or with the proteinaceous material or fragment of claims 7 to 9 or 17 to 19, or with the fusion or mixed protein of claims 36.

38. Kit comprising at least the purified antibodies or serum according to claim 37, and optionally, appropriate media or excipients for administration of the antibodies, or labelling or detection means for the antibodies.



pILL199
(35 kb)
(E, B)
Pv
Bg
Pv
E
(B*, Ps)
5 kb
pILL205
(9.5 kb)
H
(Bg)
H H
C
Ps
a
H
c
d
b
e
f
g
h
H
Ps
H
(Bg*)
i
ureA
ureB
1 kb
HindIII
PstI
1.7
0.2
2.8
2.8
1.2
0.7
3.0
4.8
1.7

- FIGURE 1 -

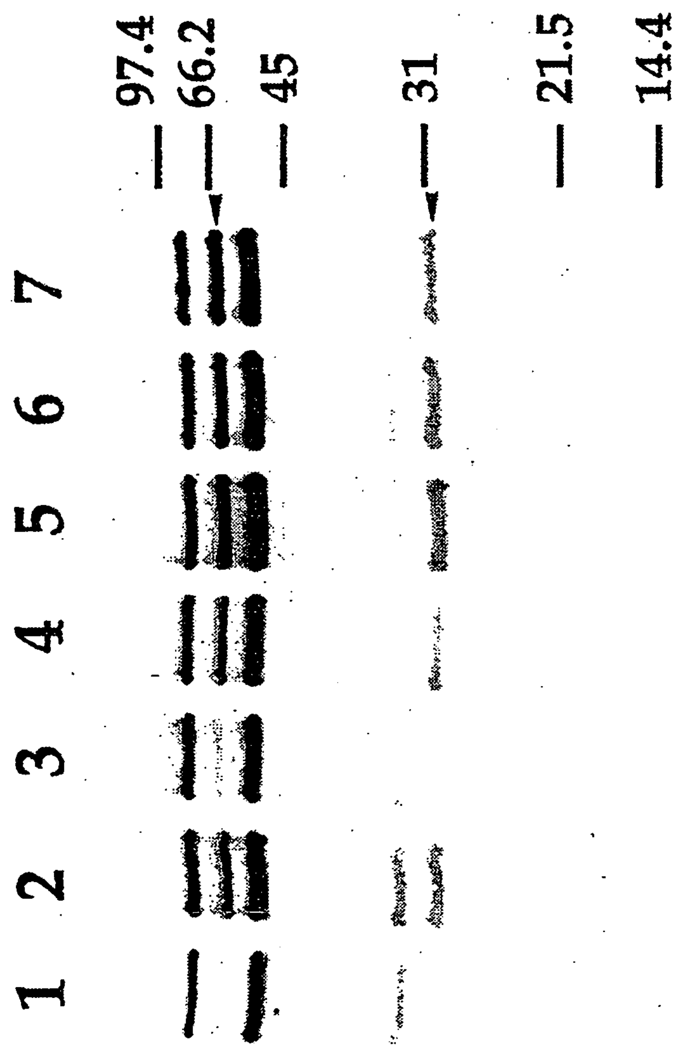


FIGURE 2 A

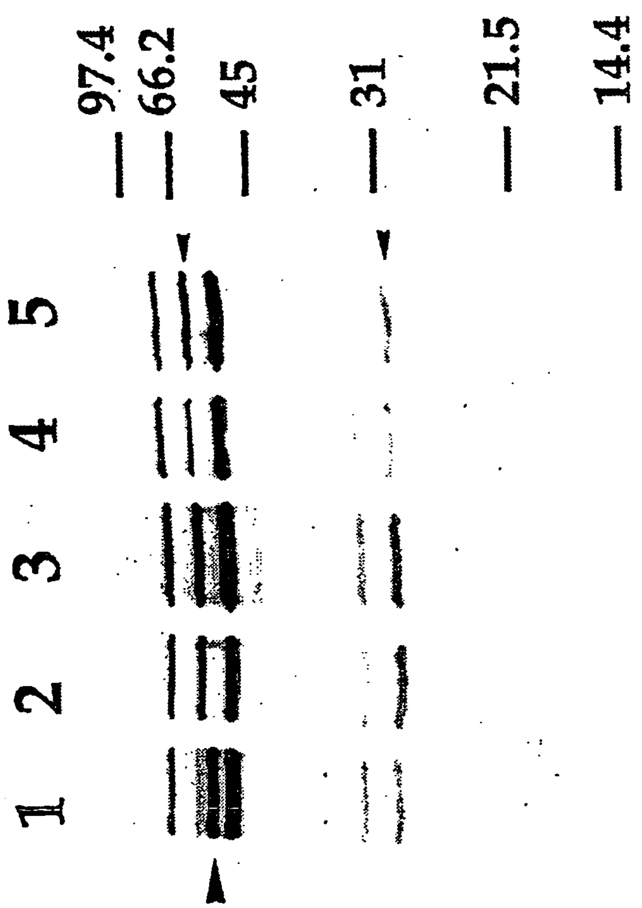


FIGURE 2 B

```
1                                           31  SD
TGA TAG CTT GGC TAC CAA TAG AAA TTC AAT AAG GAG TTT AGG ATG AAA CTA ACG CCT AAA
                                                    ureA     Met lys leu thr pro lys
61/7                                        91/17
GAA CTA GAC AAG TTA ATG CTC CAT TAT GCG GGC AGA TTG GCA GAA GAA CGC TTG GCG CGT
glu leu asp lys leu met leu his tyr ala gly arg leu ala glu glu arg leu ala arg
121/27                                      151/37
GGT GTG AAA CTC AAT TAC ACC GAA GCG GTC GCG CTC ATT AGC GGG CGT GTG ATG GAA AAG
gly val lys leu asn tyr thr glu ala val ala leu ile ser gly arg val met glu lys
181/47                                      211/57
GCG CGT GAT GGT AAT AAA AGC GTG GCG GAT TTG ATG CAA GAA GGC AGG ACT TGG CTT AAA
ala arg asp gly asn lys ser val ala asp leu met gln glu gly arg thr trp leu lys
241/67                                      271/77
AAA GAA AAT GTG ATG GAC GGC GTA GCA AGC ATG ATT CAT GAA GTG GGG ATT GAA GCT AAC
lys glu asn val met asp gly val ala ser met ile his glu val gly ile glu ala asn
301/87                                      331/97
TTC CCC GAT GGA ACC AAG CTT GTA ACT ATC CAC ACT CCG GTA GAG GAT AAT GGC AAA TTA
phe pro asp gly thr lys leu val thr ile his thr pro val glu asp asn gly lys leu
361/107                                     391/117
GCC CCC GGC GAG GTC TTC TTA AAA AAT GAG GAC ATT ACT ATT AAC GCC GGC AAA GAA GCC
ala pro gly glu val phe leu lys asn glu asp ile thr ile asn ala gly lys glu ala
```

- FIGURE 3 (1) -

```
421/127                                          451/137
ATT AGC TTG AAA GTG AAA AAT AAA GGC GAT CGT CCT GTG CAG GTG GGA TCA CAT TTC CAC
ile ser leu lys val lys asn lys gly asp arg pro val gln val gly ser his phe his
481/147                                          511/157
TTC TTC GAA GTG AAT AAG CTC TTG GAC TTC GAT CGC GCA AAA AGC TTT TGC AAA CGC CTA
phe phe glu val asn lys leu leu asp phe asp arg ala lys ser phe cys lys arg l u
541/167                                          571/177
GAC ATT GCA TCT GGA ACA GCG GTG CGC TTT GAA CCC GGG GAG GAA AAA AGT GTG GAA CTC
asp ile ala ser gly thr ala val arg phe glu pro gly glu glu lys ser val glu leu
601/187                                          631/197
ATT GAC ATC GGC GGG AAT AAG CGC ATC TAT GGC TTT AAT TCT TTG GTG GAT CGC CAA GCC
ile asp ile gly gly asn lys arg ile tyr gly phe asn ser leu val asp arg gln ala
661/207                                          691/217
GAT GCC GAT GGT AAA AAA CTC GGC TTA AAA CGC GCT AAA GAA AAA GGT TTT GGG TCT GTA
asp ala asp gly lys lys leu gly leu lys arg ala lys glu lys gly phe gly ser val
721/227                                          751     SD
AAC TGC GGT TGT GAA GCG ACT AAA GAT AAA CAA TAA GGA AAA ACC ATG AAA AAG ATT TCA
asn cys gly cys glu ala thr lys asp lys gln OCH     ureB    Met lys lys ile ser
781/6                                            811/16
CGA AAA GAA TAT GTT TCT ATG TAT GGT CCC ACT ACC GGG GAT CGT GTT AGA CTC GGC GAC
arg lys glu tyr val ser met tyr gly pro thr thr gly asp arg val arg leu gly asp
```

- FIGURE 3 (ii) -

841/26 871/36
ACT GAT TTG ATC TTA GAA GTG GAG CAT GAT TGC ACC ACT TAT GGT GAA GAG ATC AAA TTT
thr asp leu ile leu glu val glu his asp cys thr thr tyr gly glu glu ile lys phe
901/46 931/56
GGG GGC GGT AAA ACT ATC CGT GAT GGG ATG AGT CAA ACC AAT AGC CCT AGC TCT TAT GAA
gly gly gly lys thr ile arg asp gly met ser gln thr asn ser pro ser ser tyr glu

961/66 991/76
TTA GAT TTG GTG CTC ACT AAC GCC CTC ATT GTG GAC TAT ACG GGC ATT TAC AAA GCC GAC
leu asp leu val leu thr asn ala leu ile val asp tyr thr gly ile tyr lys ala asp
1021/86 1051/96
ATT GGG ATT AAA GAC GGC AAG ATT GCA GGC ATT GGC AAG GCA GGC AAT AAG GAC ATG CAA
il gly ile lys asp gly lys ile ala gly ile gly lys ala gly asn lys asp met gln
1081/106 1111/116
GAT GGC GTA GAT AAT AAT CTT TGC GTA GGT CCT GCT ACA GAG GCT TTG GCA GCT GAG GGC
asp gly val asp asn asn leu cys val gly pro ala thr glu ala leu ala ala glu gly
1141/126 1171/136
TTG ATT GTA ACC GCT GGT GGC ATC GAT ACG CAT ATT CAC TTT ATC TCT CCC CAA CAA ATC
leu ile val thr ala gly gly ile asp thr his ile his phe ile ser pro gln gln ile
1201/146 1231/156
CCT ACT GCT TTT GCC AGC GGG GTT ACA ACC ATG ATT GGA GGA GGC ACA GGA CCT GCG GAT
pro thr ala phe ala ser gly val thr thr met ile gly gly gly thr gly pro ala asp

- FIGURE 3 (iii) -

```
1261/166                                1291/176
GGC ACG AAT GCG ACC ACC ATC ACT CCC GGA CGC GCT AAT CTA AAA AGT ATG TTG CGT GCA
gly thr asn ala thr thr ile thr pro gly arg ala asn leu lys ser met leu arg ala
1321/186                                1351/196
GCC GAA GAA TAC GCC ATG AAT CTA GGC TTT TTG GCT AAG GGG AAT GTG TCT TAC GAA CCC
ala glu glu tyr ala met asn leu gly phe leu ala lys gly asn val ser tyr glu pro
1381/206                                1411/216
TCT TTA CGC GAT CAG ATT GAA GCA GGG GCG ATT GGT TTT AAA ATC CAC GAA GAC TGG GGA
ser leu arg asp gln ile glu ala gly ala ile gly phe lys ile his glu asp trp gly
1441/226                                1471/236
AGC ACA CCT GCA GCT ATT CAC CAC TGC CTC AAT GTC GCC GAT GAA TAC GAT GTG CAA GTG
ser thr pro ala ala ile his his cys leu asn val ala asp glu tyr asp val gln val
1501/246                                1531/256
GCT ATC CAC ACC GAT ACC CTT AAC GAG GCG GGC TGT GTA GAA GAC ACC CTA GAG GCG ATT
ala ile his thr asp thr leu asn glu ala gly cys val glu asp thr leu glu ala ile
1561/266                                1591/276
GCC GGG CGC ACC ATC CAT ACC TTC CAC ACT GAA GGG GCT GGG GGT GGA CAC GCT CCA GAT
ala gly arg thr ile his thr phe his thr glu gly ala gly gly gly his ala pro asp
1621/286                                1651/296
GTT ATC AAA ATG GCA GGG GAA TTT AAC ATT CTA CCC GCC TCT ACT AAC CCG ACC ATT CCT
val il  lys met ala gly glu phe asn ile leu pro ala ser thr asn pro thr ile pro
1681/306                                1711/316
TTC ACC AAA AAC ACT GAA GCC GAG CAC ATG GAC ATG TTA ATG GTG TGC CAC CAC TTG GAT
phe thr lys asn thr glu ala glu his met asp met leu met val cys his his leu asp
```

- FIGURE 3 (iv)

```
1741/326                                1771/336
AAA AGT ATC AAG GAA GAT GTG CAG TTT GCC GAT TCG AGG ATT CGC CCC CAA ACT ATC GCG
lys ser ile lys glu asp val gln phe ala asp ser arg ile arg pro gln thr ile ala
1801/346                                1831/356
GCT GAA GAC CAA CTC CAT GAC ATG GGG ATC TTT TCT ATC ACC AGC TCC GAC TCT CAG GCT
ala glu asp gln leu his asp met gly ile phe ser ile thr ser ser asp ser gln ala
1861/366                                1891/376
ATG GGA CGC GTA GGC GAG GTG ATC ACA CGC ACT TGG CAG ACA GCA GAC AAA AAC AAA AAA
met gly arg val gly glu val ile thr arg thr trp gln thr ala asp lys asn lys lys
1921/386                                1951/396
GAG TTT GGG CGC TTG AAA GAG GAA AAA GGC GAT AAC GAC AAC TTC CGC ATC AAA CGC TAC
glu phe gly arg leu lys glu glu lys gly asp asn asp asn phe arg ile lys arg tyr
1981/406                                2011/416
ATC TCT AAA TAC ACC ATC AAC CCC GGG ATC GCG CAT GGG ATT TCT GAC TAT GTG GGC TCT
ile ser lys tyr thr ile asn pro gly ile ala his gly ile ser asp tyr val gly ser
2041/426                                2071/436
GTG GAA GTG GGC AAA TAC GCC GAC CTC GTG CTT TGG AGT CCG GCT TTC TTT GGC ATT AAG
val glu val gly lys tyr ala asp leu val leu trp ser pro ala phe phe gly ile lys
2101/446                                2131/456
CCC AAT ATG ATT ATT AAG GGC GGA TTT ATT GCG CTC TCT CAA ATG GGC GAT GCC AAT GCG
pro asn met ile ile lys gly gly phe ile ala leu ser gln met gly asp ala asn ala
2161/466                                2191/476
TCT ATT CCC ACC CCT CAG CCC GTC TAT TAC CGT GAA ATG TTT GGA CAC CAT GGG AAA AAC
ser ile pro thr pro gln pro val tyr tyr arg glu met phe gly his his gly lys asn
```

- FIGURE 3 (v) -

```
2221/486                                2251/496
AAA TTC GAC ACC AAT ATC ACT TTC GTG TCC CAA GCG GCT TAC AAG GCA GGG ATC AAA GAA
lys phe asp thr asn ile thr phe val ser gln ala ala tyr lys ala gly ile lys glu
2281/506                                2311/516
GAA CTA GGG CTA GAT CGC GCG GCA CCG CCA GTG AAA AAC TGT CGC AAT ATC ACT AAA AAG
glu leu gly leu asp arg ala ala pro pro val lys asn cys arg asn ile thr lys lys
2341/526                                2371/536
GAC CTC AAA TTC AAC GAT GTG ACC GCA CAT ATT GAT GTC AAC CCT GAA ACC TAT AAG GTG
asp leu lys phe asn asp val thr ala his ile asp val asn pro glu thr tyr lys val
2401/546                                2431/556
AAA GTG GAT GGC AAA GAG GTA ACC TCT AAA GCA GCA GAT GAA TTG AGC CTA GCG CAA CTT
lys val asp gly lys glu val thr ser lys ala ala asp glu leu ser leu ala gln leu
2461/566                                2491
TAT AAT TTG TTC TAG GAG GCT AAG GAG GGG GAT AGA GGG GGT TAA TTT AGA GGG GAG TCA
tyr asn leu phe AMB
2521                                    2551
TTG ATT TAC CTT TGC TAG TTT ATA ATG GAT TTA AGA GAG GTT TTT TTT CGT GTT TTA TAC

2581                                    2611
CGC GTT GAA ACC CTC AAA TCT TTA CCA AAA GGA TGG TAA
```

- FIGURE 3 (vi) -

ureA

```
              •  •    •  •
H.f.  MKLTPKELDKLMLHYAGRLAEERLARGVKLNYTEAVALISGRVME
H.p.  *****************E**KKRKEK*I****V*******AHI**
P. m. *E***R*K***L*FT**LV**RR**K*L****P*R*****CAI**
J.b.  ***S*R*VE**G**N**Y**QKR*****R***T******ASQI**
      =  = =  == = ==  =  =   =  === = ====    ==

                 •    •    •  ••  ••  ••
      KARDGNKSVADLMQEGRTWLKKENVMDGVASMIHEVGIEANFPDG 89
      E**A*K*TA*E*******L**PDD****************M**** 89
      G**E*-*T**Q**S****V*TA*Q**E**PE**KD*QV*CT**** 89
      Y****E*T**Q**CL*QHL*GRRQ*LPA*PHLLNA*QV**TE*** 90
       == = =  = == =  =   =  =     = =  ===
```

- FIGURE 4 (i) -

```
                     •                                             • ••••
H.f.      TKLVTIHTPV---------------------------EDNGKLAPGEV
H.p.      *****V***I---------------------------*A****V***L
P. m.     ****S**S*IV 100                            1  MI***I
J.b.      *****V*D*ISRENGELQEALFGSLLPVPSLDKFAETKEDNRI***I
          ====  = =                                     ===


        •••• ••        •     •     •               •  •
        FLKNEDITI--NAGKEAISLKVKNKGDRPVQVGSHFHFFEVNKLL 154
        *********--*E**K*V*V****V******I***********RC* 154
        RVNAALGD*EL***R*TKTIQ*A*H******C***Y**Y***EA* 51
        LCED*CL*L--*I*RK*VI***TS*****I*****Y**I***PY* 180
                  = =      =    ==== = === == ===  =
```

- FIGURE 4 (11) -

```
           •       • •  •  •            ••      •• •  •
H.f.     DFDRAKSFCKRLDIASGTAVRFEPGEEKSV-ELIDIGGNKRIY
H.p.     ****E*T*G*********************-********R**F
P. m.    R*A*KETLGF**N*PA*M*******QSRT*D**VAFA*KRE**
J.b.     T***R*AYGM**N**AG********DC***-T*VS*E**KV*R
          = =      == =    =======        =    =  =

 • •• ••••  •• • ••••
GFNSLVDRQADADGKKLGLKRAKEKGF-GSVNCGCEAT---KDKQ      237
***A*******NES**IA*H****R**H*AKSDDNYVKTI-*E        238
--------------------------**H*KVMGKLESE--K*        109
*G*AIA*GPVNETNLEAAMHAVRSR**-*HEEEKDAPEGFT*EDPNCSF-270
                          == =        =
```

- FIGURE 4 (iii) -

ureB

```
             •      ••         •      •  •   •  •
H.f.   MKKISRKEYVSMYGPTTGDRVRLGDTDLILEVEHDC
H.p.   ********************K********A*****Y
P. m.  **T***QA*AD*F*******L**A**E*F**I*K*F
J.b.   NTF*H****ANK*******KI*****N*LA*I*K*Y
          = =  =    ======  == ==  =  = = =

             •     •  •  ••  •       ••
TTYGFEIKFGGGKTIRDGMSQTNSPSSYEL-DLVLTNALIVDYTGI   81
*I****L*******L*E****S*N**KE**-**II***********   81
*****V******V*****G*SQVV*AECV-*VLI***I*L**W**   81
AL**D*CV*****V*****G*SCGHPPAISL*T*I***V*I*****   352
  == =  =====  = == =            =  === = == ==
```

- FIGURE 4 (iv) -

```
          •         •         • • •   • • •  •  ••
H.f.  YKADIGIKDGKIAGIGKAGNKDMQDGVDNNLCVGPATEALAAEGL
H.p.  *****************G*********K***S*********G***
P. m. V*********R*V*******P*VQPN**IVI--**G**VV*G**K
J.b.  I*********L**S******P*IMN**FSNMII*AN**VI*G***
       ========= =  === == =    =      =  ==  = ==

               •   ••• ••       •        •
      IVTAGGIDTHIHFISPQQIPTAFASGVTTMIGGGTGPADGTNATT 171
      ********************************************* 171
      **************C***AQEGLV*****FI******VA****** 169
      ********C***Y*C**LVYEAIS**I**LV*******A**R*** 442
      ======== === = ==      == ==  ======  == ===
```

- FIGURE 4 (v) -

```
         •   •  ••  •• • • • ••  •    •  •  •
H.f.     ITPGRANLKSMLRAAEEYAMNLGFLAKGNVSYEPSLRDQIEAGAI
H.p.     *****R***W********S**********A*NDA*A*********
P. m.    V***IW*MYR**E*VD*LPI*V*LFG**CV*QPEAI*E**T****
J.b.     C**SPTQMRL**QSTDDLPL*F**TG**SS*KPDE*HEI*K***M
          ==      ==        = =   ==  =        ===

 •              •       •              • •
GFKIHEDWGSTPAAIHHCLNVADEYDVQVAIHTDTLNEAGCVEDT 261
*********T**S**N*A*D***K********************* 261
*L*******A**M***N*******M*******S*****G*FY*E* 259
*L*L***********DN**TI*EHH*I*IN**********F**HS 532
= = ===== == ==   =  =   = =  == =====  =  =
```

- FIGURE 4 (vi) -

```
                                         •           • •
H.f.     LEAIAGRTIHTFHTEGAGGGHAPDVIKMAGEFNILPASTNPTIPF
H.p.     MA*****TM****************I**V***H************
P. m.    VK*****VI*****************SV**P***********M*Y
J.b.     IA*FK**TI**Y*S***********I**VC*IK*V**S*****R*L
          =  ==  == = ========== ==  =  = == ===== =


   ••  •         •   •    •  •    ••
TKNTEAEHMDMLMVCHHLDKSIKEDVQFADSRIRPQTIAAEDQLH  351
*V*****************************************T**  351
*I**VD**L**********P**P***A**F****RE*******I**  349
*S**ID**L**********RE*P***A**H****KK*******V**  622
= ==  == ==========  = === == ====  ====== ==
```

- FIGURE 4 (vii) -

```
                 ••  •           •            ••• •  ••••
H.f.     DMGIFSITSSDSQAMGRVGEVITRTWQTADKNKKEFGRLKEEKGD
H.p.     *********************************************
P. m.    ***AI*VM**************L****C*H***LQR*T*AGDSA*
J.b.     *I*AI**I**************S**********AQT*P**CDSS*
         = =  =  ============= ==== = ===   = =

            •                           •
NDNFRIKRYISKYTINPGIAHGISDYVGSVEVGKYADLVLWSPAF 441
*********L*******A******E*********V********** 441
***N******A******AL****AHT***I*K**L**I***D*** 439
**********A******A******Q*********L******K*S* 712
=== =====  ======  ====   === = == == === = =
```

- FIGURE 4 (viii) -

```
H.f.   FGIKPNMIIKGGFIALSQMGDANASIPTPQPVYYREMFGHHGKNK
H.p.   **V**********************************A*****A*
P. m.  **V**AL*****MVRYAP***I**A********H**P*YACL**A*
J.b.   **T**E*V****MVAWADI**P**********KM*P*Y*TL**AG

FDTNITFVSQAAYKAGIKEELGLDRAAPPVKN--CRNITKKDLKF 529
Y*R**********DK*********E*QVL****--*********MQ* 529
YQ*SMI*M*K*GIEA*VP*K***KSLSLIGRVEGC*H***ASMIH 529
GALS*A***K**LDQRVNVLY**NKRVEA*S*--V*KL**L*M*L 800
```

- FIGURE 4 (ix) -

```
        ••    •       •    •••••       •   •
H.f.   NDVTAHIDVNPETYKVKVDGKEVTSKAADELSLAQLYNLF  569
H.p.   **T****E******H*F*********P*NKV*****FSI*  569
P. m.  *NYVP**ELD*Q**I**A**VPLVCEP*T**PM**R*F**  569
J.b.   **ALPE*T*D**S*T**A***LLCVSE*TTVP*SRN*F**  840
       =     =   =  = =  ==       =          =
```

ureA	:74 % identity	*ureB*	:88 % identity
ureA	:46 % identity	*ureB*	:62 % identity
ureA	:47 % identity	*ureB*	:59 % identity

- FIGURE 4 (x) -

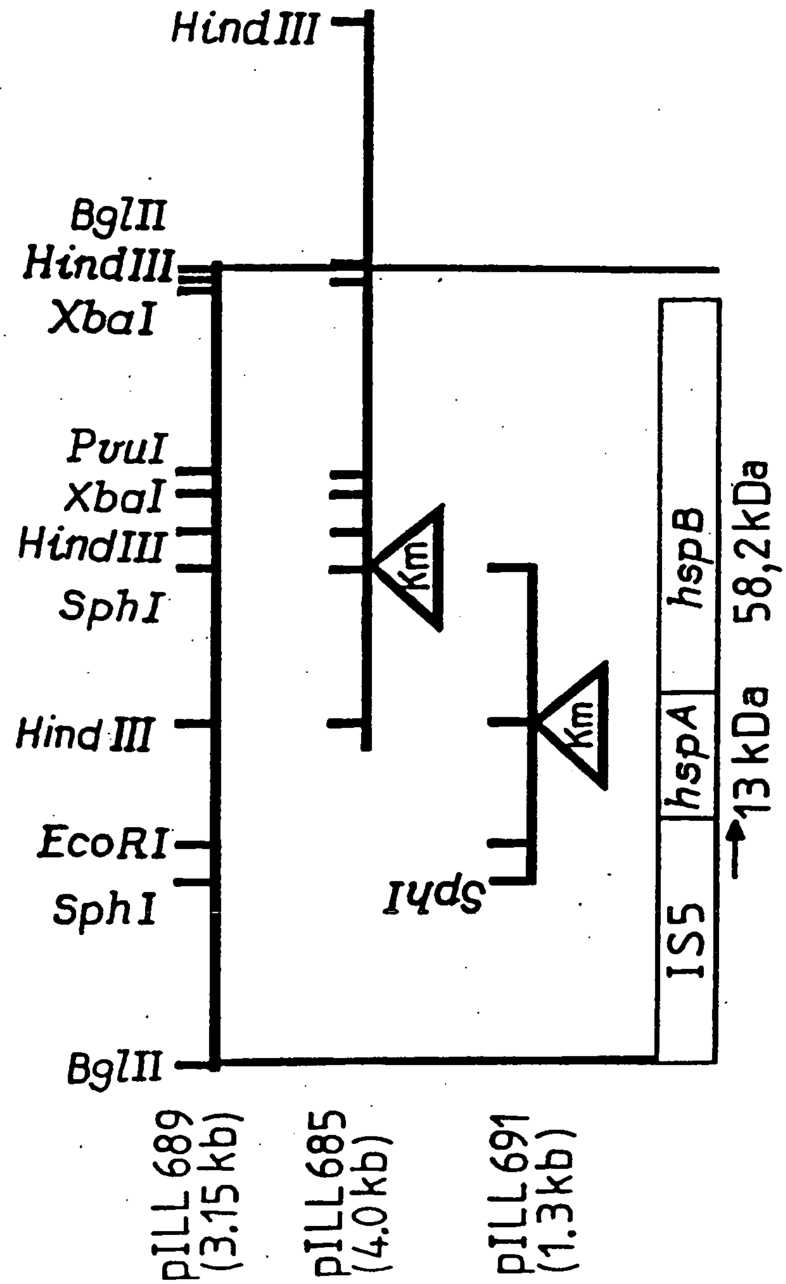

HindIII
BglII
HindIII
XbaI
PvuI
XbaI
HindIII
SphI
Hind III
EcoRI
SphI
BglII
Km
Km
SphI
IS5
hspA
hspB
13 kDa
58,2kDa
pILL 689 (3.15 kb)
pILL 685 (4.0 kb)
pILL 691 (1.3 kb)

FIGURE 5

```
1                                       31
ACA AAC ATG ATC TCA TAT CAG GGA CTT GTT CGC ACC TTC CCT AAA AAT GCG CTA TAG TTG

61                                      91
TGT CGC TTA AGA ATA CTA AGC GCT AAA TTT CTA TTT TAT TTA TCA AAA CTT AGG AGA ACT

121                                     151/10
GAA ATG AAG TTT CAA CCA TTA GGA GAA AGG GTC TTA GTA GAA AGA CTT GAA GAA GAG AAC
    met lys phe gln pro leu gly glu arg val leu val glu arg leu glu glu glu asn

181/21                                  211/31
AAA ACC AGT TCA GGC ATC ATC ATC CCT GAT AAC GCT AAA GAA AAG CCT TTA ATG GGC GTA
lys thr ser ser gly ile ile ile pro asp asn ala lys glu lys pro leu met gly val

241/41                                  271/51
GTC AAA GCG GTT AGC CAT AAA ATC AGT GAG GGT TGC AAA TGC GTT AAA GAA GGC GAT GTG
val lys ala val ser his lys ile ser glu gly cys lys cys val lys glu gly asp val

301/51                                  331/71
ATC GCT TTT GGC AAA TAC AAA GGC GCA GAA ATC GTT TTA GAT GGC GTT GAA TAC ATG GTG
ile ala phe gly lys tyr lys gly ala glu ile val leu asp gly val glu tyr met val

361/71                                  391/91
CTA GAA CTA GAA GAC ATT CTA GGT ATT GTG GGC TCA GGC TCT TGC TGT CAT ACA GGT AAT
leu glu leu glu asp ile leu gly ile val gly ser gly ser cys cys his thr gly asn

421/91                                  451/111
CAT GAT CAT AAA CAT GCT AAA GAG CAT GAA GCT TGC TGT CAT GAT CAC AAA AAA CAC TAA
his asp his lys his ala lys glu his glu ala cys cys his asp his lys lys his OCH

481                                     511
AAA ACA TTA TTA TTA AGG ATA CAA AAT GGC AAA AGA
```

- FIGURE 6 (1) -

```
479                                     509/2
AAA AAA CAT TAT TAT TAA GGA TAC AAA ATG GCA AAA GAA ATC AAA TTT TCA GAT AGC GCA
                                    met ala lys glu ile lys phe ser asp ser ala

539/12                                  569/22
AGA AAC CTT TTA TTT GAA GGC GTA AGA CAA CTC CAT GAC GCT GTC AAA GTA ACC ATG GGG
arg asn leu leu phe glu gly val arg gln leu his asp ala val lys val thr met gly

599/32                                  629/42
CCA AGA GGC AGG AAC GTG TTG ATC CAA AAA AGC TAT GGC GCT CCA AGC ATC ACC AAA GAC
pro arg gly arg asn val leu ile gln lys ser tyr gly ala pro ser ile thr lys asp

659/52                                  689/62
GGC GTG AGC GTG GCT AAA GAG ATT GAA TTA AGT TGC CCC GTG GCT AAC ATG GGC GCT CAG
gly val ser val ala lys glu ile glu leu ser cys pro val ala asn met gly ala gln

719/72                                  749/82
CTC GTT AAA GAA GAT GCG AGC AAA ACC GCT GAT GCC GCC GGC GAT GGC ACG ACC ACA GCG
leu val lys glu asp ala ser lys thr ala asp ala ala gly asp gly thr thr thr ala

779/92                                  809/102
ACC GTG CTG GCT TAT AGC ATT TTT AAA GAG GGC TTG AGG AAT ATC ACG GCT GGG GCT AAC
thr val leu ala tyr ser ile phe lys glu gly leu arg asn ile thr ala gly ala asn

839/112                                 869/122
CCT ATT GAA GTG AAA CGA GGC ATG GAT AAA GCG CCT GAA GCG ATC ATT AAT GAG CTT AAA
pro ile glu val lys arg gly met asp lys ala pro glu ala ile ile asn glu leu lys

899/132                                 929/142
AAA GCG AGC AAA AAA GTG GGC GGT AAA GAA GAA ATC ACC CAA GTA GCG ACC ATT TCT GCA
lys ala ser lys lys val gly gly lys glu glu ile thr gln val ala thr ile ser ala
```

- FIGURE 6 (11) -

```
959/152                                         989/162
AAC TCC GAT CAC AAT ATC GGG AAA CTC ATC GCT GAC GCT ATG GAA AAA GTG GGT AAA GAC
asn s r asp his asn ile gly lys leu ile ala asp ala met glu lys val gly lys asp

1019/172                                        1049/182
GGC GTG ATC ACC GTT GAA GAA GCT AAG GGC ATT GAA GAT GAA TTA GAT GTC GTA GAA GGC
gly val ile thr val glu glu ala lys gly ile glu asp glu leu asp val val glu gly

1079/192                                        1109/202
ATG CAA TTT GAT AGA GGC TAC CTC TCC CCT TAC TTT GTA ACC AAC GCT GAG AAA ATG ACC
met gln phe asp arg gly tyr leu ser pro tyr phe val thr asn ala glu lys met thr

1139/212                                        1169/222
GCT CAA TTG GAT AAC GCT TAC ATC CTT TTA ACG GAT AAA AAA ATC TCT AGC ATG AAA GAC
ala gln leu asp asn ala tyr ile leu leu thr asp lys lys ile ser ser met lys asp

1199/232                                        1229/242
ATT CTC CCG CTA CTA GAA AAA ACC ATG AAA GAG GGC AAA CCG CTT TTA ATC ATC GCT GAA
il  l u pro leu leu glu lys thr met lys glu gly lys pro leu leu ile ile ala glu

1259/252                                        1289/262
GAC ATT GAG GGC GAA GCT TTA ACG ACT CTA GTG GTG AAT AAA TTA AGA GGC GTG TTG AAT
asp ile glu gly glu ala leu thr thr leu val val asn lys leu arg gly val leu asn

1319/272                                        1349/282
ATC GCA GCG GTT AAA GCT CCA GGC TTT GGG GAC AGG AGA AAA GAA ATG CTC AAA GAC ATC
ile ala ala val lys ala pro gly phe gly asp arg arg lys glu met leu lys asp ile
```

- FIGURE 6 (iii) -

```
1379/292                                1409/302
GCT GTT TTA ACC GGC GGT CAA GTC ATT AGC GAA GAA TTG GGC TTG AGT CTA GAA AAC GCT
ala val leu thr gly gly gln val ile ser glu glu leu gly leu ser leu glu asn ala

1439/312                                1469/322
GAA GTG GAG TTT TTA GGC AAA GCG AAG ATT GTG ATT GAC AAA GAC AAC ACC ACG ATC GTA
glu val glu phe leu gly lys ala lys ile val ile asp lys asp asn thr thr ile val

1499/332                                1529/342
GAT GGC AAA GGC CAT AGC CAT GAC GTC AAA GAC AGA GTC GCG CAA ATC AAA ACC CAA ATT
asp gly lys gly his ser his asp val lys asp arg val ala gln ile lys thr gln ile

1559/352                                1589/362
GCA AGC ACG ACA AGC GAT TAC GAC AAA GAA AAA TTG CAA GAA AGA TTG GCC AAA CTC TCT
ala ser thr thr ser asp tyr asp lys glu lys leu gln glu arg leu ala lys leu ser

1619/372                                1649/382
GGC GGT GTG GCT GTG ATT AAA GTG GGC GCT GCG AGT GAA GTG GAA ATG AAA GAG AAA AAA
gly gly val ala val ile lys val gly ala ala ser glu val glu met lys glu lys lys

1679/392                                1709/402
GAC CGG GTG GAT GAC GCG TTG AGC GCG ACT AAA GCG GCG GTT GAA GAA GGC ATT GTG ATT
asp arg val asp asp ala leu ser ala thr lys ala ala val glu glu gly ile val ile

1739/412                                1769/422
GGG GGC GGT GCG GCC CTC ATT CGC GCG GCC CAA AAA GTG CAT TTG AAT TTA CAC GAT GAT
gly gly gly ala ala leu ile arg ala ala gln lys val his leu asn leu his asp asp
```

- FIGURE 6 (iv) -

```
1799/432                                1829/442
GAA AAA GTG GGC TAT GAA ATC ATC ATG CGC GCC ATT AAA GCC CCA TTA GCT CAA ATC GCT
glu lys val gly tyr glu ile ile met arg ala ile lys ala pro leu ala gln ile ala

1859/452                                1889/462
ATC AAT GCC GGT TAT GAT GGC GGT GTG GTC GTG AAT GAA GTA GAA AAA CAC GAA GGG CAT
ile asn ala gly tyr asp gly gly val val val asn glu val glu lys his glu gly his

1919/472                                1949/482
TTT GGT TTT AAC GCT AGC AAT GGC AAG TAT GTG GAC ATG TTT AAA GAA GGC ATT ATT GAC
phe gly phe asn ala ser asn gly lys tyr val asp met phe lys glu gly ile ile asp

1979/492                                2009/502
CCC TTA AAA GTA GAA AGG ATC GCT TTA CAA AAT GCG GTT TCG GTT TCA AGC CTG CTT TTA
pro leu lys val glu arg ile ala leu gln asn ala val ser val ser ser leu leu l u

2039/512                                2069/522
ACC ACA GAA GCC ACC GTG CAT GAA ATC AAA GAA GAA AAA GCG GCC CCA GCA ATG CCT GAT
thr thr glu ala thr val his glu ile lys glu glu lys ala ala pro ala met pro asp

2099/532                                2129/542
ATG GGT GGC ATG GGC GGA ATG GGA GGC ATG GGC GGC ATG ATG TAA GCC CCC TTG CTT TTT
met gly gly met gly gly met gly gly met gly gly met met OCH

2159                                    2189
GGT ATC ATC TGC TTT TAA AAT CCA TCT TCT AGA ATC CCC CCT TCT AAA ATC CCT TTT TTG

2219                                    2249
GGG GGT GCT TTT GGT TTG ATA AAA CCG CTC GCT TTT AAA AAC GCG CAA CAA AAA ACT CTG

2279
TTA AGC
```

- FIGURE 6 (v) -

```
MAKEIKFSDSARNLLFEGVRQLHDAVKVTMGPRGRNVLIQKSYG
****LR*G*D**LQMLA**NA*A***Q**********VLE****
MA**DV**GND**VKMLR**NV*A******L**K****VLD**F*
MA**N**YNED**KKIHK**KT*AE*****L**K**H*V*D**F*
***T*AYDEE**RG*ER*LNS*A******L**K****VLE*KW*
Y**DV-**GAD**ALMLQ**DL*A***A*****K**T*I*EQ*W*
```

```
APSITKDGVSVAKEIELSCPVANMGAQLVKEDASKTADAAGDG
**TV************FEHRFM*****M***V****S*T****
**T*********R****EDKFE*****M***V***AN******
S*QV*****T*******EDKHE*****M***V*******K****
**T**N****I******ED*YEKI**E****V*K**D*V****
S*KV*****T***S*D*KDKYK*I**K**QDV*NN*NEE****
```

- FIGURE 7 A (I) -

```
TTTATVLAYSIFKEGLRNITAGANPIEVKRGMDKAPEAIINELKK
********R**LV**HKAVA**M**MDL***I***VL*VTKK*QA
********QA*IT***KAVA**M**MDL***I***VT*AVE***A
********EA*YS*****V******MLD***I***VKVVVD*I**
********QALV******VA*****LGL***IE**VDKVTET*L*
********R**A***FEK*SK****VEIR**V*L*VD*V*A****
```

```
ASKKVGGKEEITQVATISANSDHNIGKLIADAMEKVGKDGVIT
M**PCKDSKA*A**G*******EA**AI**E*******E****
L*VPCSDSKA*A**G*******ETV*****E**D****E****
I**P*QHHK**A********N*AE**N***E*******N*S**
DA*E*ET**Q*AAT*A***-G*QS**D***E**D***NE****
Q**P*TTP***A********G*KE**NI*SD**K***RK****
```

- FIGURE 7 A (ii) -

```
VEEAKGIEDELDVVEGMQFDRGYLSPYFVTNAEKMTAQLDNAYIL
**DGN*L*N**Y***********I****IN*QQN*SCE*EHPF**
**DGT*LQ********************INKP*TGAVE*ESPF**
******F*TV*******N*N*****S**S**P*TQECV*EE*LV*
***SNTFGLQ*ELT***R**K**I*G****D**RQE*V*EEP***
*KDG*TLN***EII***K*****I****INTSKGQKCEFQD**V*
```

```
LTDKKISSMKDILPLLEKTMKEGKPLLIIAEDIEGEALTTLVV
*V***V**IREM*SV**GVA*S*R********I*****A****
*A*****NIREM**V**AVA*A**********V*****A*A**
IY*****GI**F**V*QQVAES*R**************A****
*VSS*V*TV**L******VIQA**S*******V*****S****
*SE*****IQS*V*A**IANLVLNR*KVGLQVVAVK*PGF**L
```

- FIGURE 7 A (III) -

```
NKLRGVLNIAAVKAPGFGDRRKEMLKDIAVLTGGQVISEELGLSL
*NM**IVKVC************A**Q***I**K********I*K**
*TI**IVKV*************A**Q***T****T******I*ME*
*R**AGFRVC************A**E***I*****L*******MK*
**I**TFKSV************A**Q*M*I***A*******V**T*
*R*KVG*QVV*V********N**NQ*K*M*IA***A*FG**GLTLN
```

```
ENAEVEF-LGKAKI-VIDKDNTTIVDGKGHSHDVKDRVAQIKT
*G*TL*D-**S**RI*VT*E****I**E*KATEINA*I***RA
*K*TL*D-**Q**RV**N**T***I**V*EEAAIQG*****RQ
**TTLAM-*****KVIVS*ED****E*L*SKE*IES*CES**K
**TDLSL-****RKV*MT**E****E*A*DTDAIAG*****R*
LEDVQPHD***VGEVIVT**DAMLLK*K*DKAQIEK*IQE*IE
```

- FIGURE 7 A (iv) -

```
QIASTTSDYDKEKLQERLAKLSGGVAVIKVGAASEVEMKEKKD
*MEE******R******V***A***********T********A
**EEA*****R******V***A***********T********A
**EDS***********************R****T*I*******
E*ENSD****R**********A*******A***T***L**R*H
*LDV***E*EK***N******SD****L***GT*D**VN****
```

```
RVDDALSATKAAVEEGIVIGGGAALIRAAQKVH---LN-LHDDEK
**E***H**R********A***V*****QKALDS--*KGDN**QN
**E***H**R******V*A***V****V*S*LAD--*RGQNE*QN
*****QHA*L*******LP***T**V*CIPTLEAFIPILTNE**Q
*IE**VRNA*********A***VT*LQ**PALDK--*K-*TG**A
**T***N**R********L***C**L*CIPALDS--*TPANE*Q*
```

- FIGURE 7 A (v) -

```
VGYEIIMRAIKAPLAQIAINAGYDGGVVVNEVEKHEGHFGFNA
M*IN*LR***ES*MR**VT****EAS****K*AE*KDNY****
**IKVAL**ME***R**VL*C*EEPS**A*T*KGGD*NY*Y**
I*AR*VLK*LS***K***A***KE*AIICQQ*LSRSSSE*YD*
T*AN*VKV*LE***K***F*S*MEP***AEK*RNLSVGH*L**
I*I***K*TL*I*AMT**K***V**SLI*EKIMQSSSEVGYD*
```

```
SNGKYVDMFKEGIIDPLKVERIALQNAVSVSSLLLTTEATVHEIK
AT*E*G**VEM**L**T**T*M*****A**A**M****CM*ADLP
ATEE*GN*IDM**L**T**T*S***Y*A**AG*MI***CM*TDLP
LRDA*T**IEA**L**T**T*C**ES*A**AG*******LIAD*P
AT*EYE*LL*A*VA**V**T*S*****A*IAG*F*****V*ADKP
MA*DF*N*VEK*****T**V*T**LD*A**A***T*A*VV*T**P
```

- FIGURE 7 A (vi) -

EEKAAPAMPDMGGMGGMGGMGGMM	**HspB** ***Helicobacter pylori***
KKEEGVGAG***************	**HtpB** ***Legionella pneumophila***
KND**-DLGAA*************	**GroEL** ***Escherichia coli***
***SSSA-*A*P*A*-*DY	**HypB** ***Chlamydia psittaci***
*KT****SDPTGGMGGMDF	**GroEL1** ***Mycobacterium leprae***
D-*G*GA***-*-M**G*F	**63 kDa Human mitochondrial protein P1**

Identity : 62.7%
60.5%
59.6%
57.4%
42.5%

Comparison of the GroEL-like proteins from various bacteria

- FIGURE 7 A (vii) -

```
       Helicobacter pylori          MKFQPLGERVL
35%    Mycobacterium leprae              **EDKI*
35.6%  Legionella pneumophila       **IR**HD**V
33.8%  Thermophilic bacterium       *LK-***D*IV
32.2%  Clostridium perfringens      *SIK***D**V
20.3%  Escherichia coli             MNIR**HD**I

       VERLEEENKTSSGIIIPDNAKEKPLMGVVKAV---SHKI
       *QAG*A*TM*P**LV**ED*****QE*T*V**GPGRWDE
       *R*M***RT*AG**V***S*T***MR*EII**GAGKVLE
       I*VV*T****A***VL**T*****QE*R*V**GAGRVLD
       IK***A*ET*K****VTGT***R*QEAE*V**GPGAIVD
       *K*K*V*T*SAG**VLTGS*AA*STR*E*L**GNGRILE
```

- FIGURE 7 B (1) -

```
SEGCKC---VKEGDVIAFGKYKGAEIVLDGVEYMVLELE
DGAKRIPVD*S***IVIYS**G*T**KYN*E**LI*SAR
NGDVRA---**V***VL****S*T*V*V**K*LV*MRED
NGQRIGRKS-*V**RVI*S**A*T*VKY**K*Y*I*RES
-GKRTEME-**I**KVLYS**A*T*VKFE*E**TI*RQD
NGEVKP-LD**VG*IVI*NDGY*VKSEKIDN*EVLIMS*

DILGIVGSGSCCHTGNHDHKHAKEHEACCHDHKKH
*V*AV*SK
**M*VIEK
***AVIR
***A**E
SDILAIVEA
```

Comparison of the GroES-like proteins from various bacteria

- FIGURE 7 B (ii) -

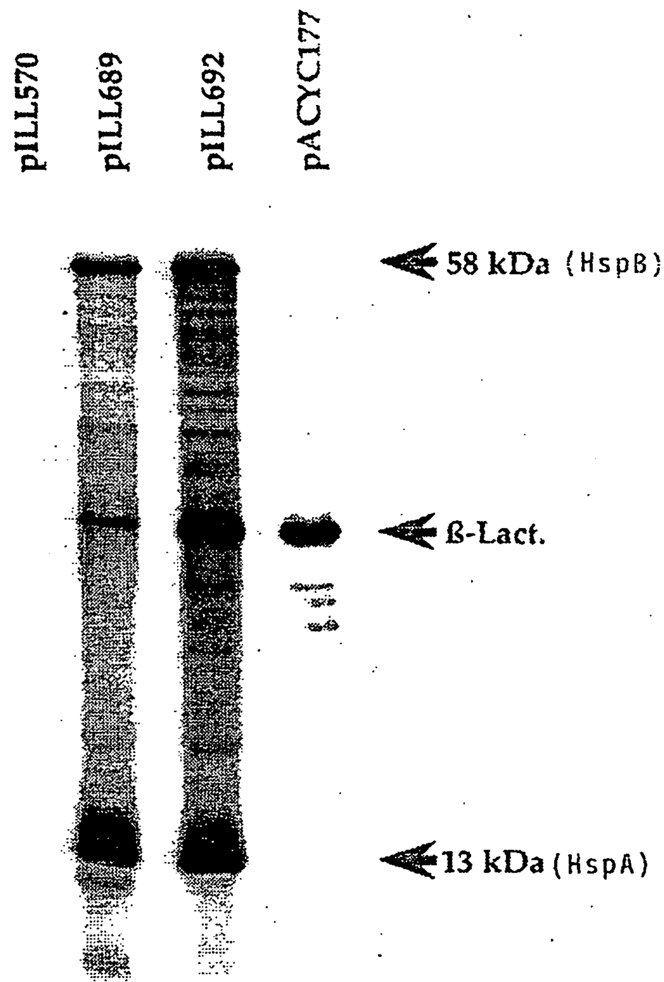

pILL570
pILL689
pILL692
pACYC177
58 kDa (HspB)
ß-Lact.
13 kDa (HspA)

FIGURE 8

1/1
ATG TTA GGT CTT GTG TTA TTG TAT GTT GCG
Met leu gly leu val leu leu tyr val ala

31/11
GTC GTG CTG ATC AGC AAC GGA GTT AGT GGG
val val leu ile ser asn gly val ser gly

- FIGURE 9 (1) -

61/21
CTT GCA AAT GTG GAT GCC AAA AGC AAA GCC
leu ala asn val asp ala lys ser lys ala

91/31
ATC ATG AAC TAC TTT GTG GGG GGG GAC TCT
ile met asn tyr phe val gly gly asp ser

- FIGURE 9 (II) -

121/41
CCA TTG TGT GTA ATG TGG TCG CTA TCA TCT
pro leu cys val met trp ser leu ser ser

151/51
TAT TCC ACT TTC CAC CCC ACC CCC CCT GCA
tyr ser thr phe his pro thr pro pro ala

- FIGURE 9 (iii) -

181/61
ACT GGT CCA GAA GAT GTC GCG CAG GTG TCT
thr gly pro glu asp val ala gln val ser

211/71
CAA CAC CTC ATT AAC TTC TAT GGT CCA GCG
gln his leu ile asn phe tyr gly pro ala

- FIGURE 9 (iv) -

241/81
ACT GGT CTA TTG TTT GGT TTT ACC TAC TTG
thr gly leu leu phe gly phe thr tyr leu

271/91
TAT GCT GCC ATC AAC AAC ACT TTC AAT CTC
tyr ala ala ile asn asn thr phe asn leu

- FIGURE 9 (v) -

301/101
GAT TGG AAA CCC TAT GGC TGG TAT TGC TTG
asp trp lys pro tyr gly trp tyr cys leu

331/111
TTT GTA ACC ATC AAC ACT ATC CCA GCG GCC
phe val thr ile asn thr ile pro ala ala

- FIGURE 9 (vi) -

361/121
ATT CTT TCT CAC TAT TCC GAT GCG CTT GAT
ile leu ser his tyr ser asp ala leu asp

391/131
GAT CAC CGC CTC TTA GGA ATC ACT GAG GGC
asp his arg leu leu gly ile thr glu gly

- FIGURE 9 (vii) -

421/141
GAT TGG TGG GCT TTC ATT TGG CTT GCT TGG
asp trp trp ala phe ile trp leu ala trp

451/151
GGT GTT TTG TGG CTC ACT GGT TGG ATT GAA
gly val leu trp leu thr gly trp ile glu

- FIGURE 9 (viii) -

481/161
TGC GCA CTT GGT AAG AGT CTA GGT AAA TTT
cys ala leu gly lys ser leu gly lys phe

511/171
GTT CCA TGG CTT GCC ATC GTC GAG GGC GTG
val pro trp leu ala ile val glu gly val

- FIGURE 9 (ix) -

541/181
ATC ACC GCT TGG ATT CCT GCT TGG CTA CTC
ile thr ala trp ile pro ala trp leu leu

571/191
TTT ATC CAA CAC TGG TCT TGA
phe ile gln his trp ser OPA

- FIGURE 9 (x) -

Comparison of the amino acid sequence of the UreI proteins deduced from the nucleotide sequence of the *ureI* gene of *H. felis* and that of *H. pylori*

Percent Similarity : 88.2
Percent Identity : 73.8

First line : *H. felis* UreI
Second line : *H. pylori* UreI

```
                .         .         .         .         .
  1 KGWMLGLVLLYVAVVLISNGVSGLANVDAKSKAIMNYFVGGDSPLCVMWS  50
       |||||||||::|||||||::||..||:||.|:||:||||  |  :|  :
  1 ...MLGLVLLYVGIVLISNGICGLTKVDPKSTAVMNFFVGGLSIICNV.V  46
                .         .         .         .         .
 51 LSSYSTFHPTPPATGPEDVAQVSQHLINFYGPATGLLFGFTYLYAAINNT 100
    : .||.::||:|..|:||:||||:||.||||||||||||||||||||||:|
 47 VITYSALNPTAPVEGAEDIAQVSHHLTNFYGPATGLLFGFTYLYAAINHT  96
                .         .         .         .         .
101 FNLDWKPYGWYCLFVTINTIPAAILSHYSDALDDHRLLGITEGDWWAFIW 150
    |.|||:||:||:|||.||||||||||||||| ||||::||||||||||:||
 97 FGLDWRPYSWYSLFVAINTIPAAILSHYSDMLDDHKVLGITEGDWWAIIW 146
                .         .         .         .
151 LAWGVLWLTGWIECALGKSLGKFVPWLAIVEGVITAWIPAWLLFIQHWS 199
    ||||||||||::||    |    .||||.|||||:||::||||||||||||||
147 LAWGVLWLTAFIENILKIPLGKFTPWLAIIEGILTAWIPAWLLFIQHWV 195
```

- FIGURE 10 -

The Genetic Code

Second Position

First Position (5' End)	U		C		A		G		Third Position (3' End)
U	UUU	Phe	UCU	Ser	UAU	Tyr	UGU	Cys	U
	UUC		UCC		UAC		UGC		C
	UUA	Leu	UCA		UAA*	*Stop*	UGA*	*Stop*	A
	UUG		UCG		UAG*	*Stop*	UGG	Trp	G
C	CUU	Leu	CCU	Pro	CAU	His	CGU	Arg	U
	CUC		CCC		CAC		CGC		C
	CUA		CCA		CAA	Gln	CGA		A
	CUG		CCG		CAG		CGG		G
A	AUU	Ile	ACU	Thr	AAU	Asn	AGU	Ser	U
	AUC		ACC		AAC		AGC		C
	AUA		ACA		AAA	Lys	AGA	Arg	A
	AUG+	Met	ACG		AAG		AGG		G
G	GUU	Val	GCU	Ala	GAU	Asp	GGU	Gly	U
	GUC		GCC		GAC		GGC		C
	GUA		GCA		GAA	Glu	GGA		A
	GUG+		GCG		GAG		GGG		G

- FIGURE 11 -

Abbreviations for amino acids

Amino acid	*Three-letter abbreviation*	*One-letter symbol*
Alanine	Ala	A
Arginine	Arg	R
Asparagine	Asn	N
Aspartic acid	Asp	D
Asparagine or aspartic acid	Asx	B
Cysteine	Cys	C
Glutamine	Gln	Q
Glutamic acid	Glu	E
Glutamine or glutamic acid	Glx	Z
Glycine	Gly	G
Histidine	His	H
Isoleucine	Ile	I
Leucine	Leu	L
Lysine	Lys	K
Methionine	Met	M
Phenylalanine	Phe	F
Proline	Pro	P
Serine	Ser	S
Threonine	Thr	T
Tryptophan	Trp	W
Tyrosine	Tyr	Y
Valine	Val	V

- FIGURE 12 -

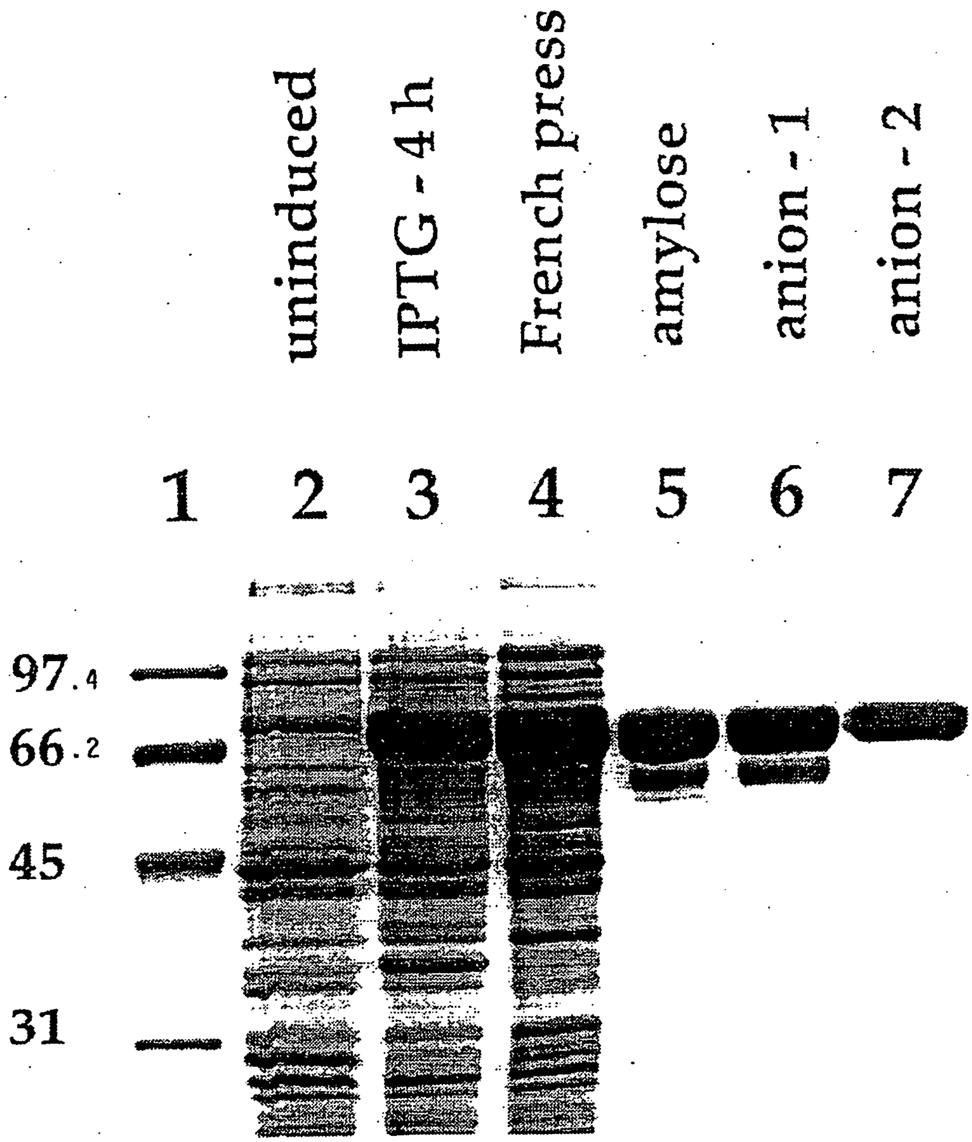

uninduced
IPTG - 4 h
French press
amylose
anion - 1
anion - 2
1
2
3
4
5
6
7
97.4
66.2
45
31

FIGURE 13

FIGURE 14

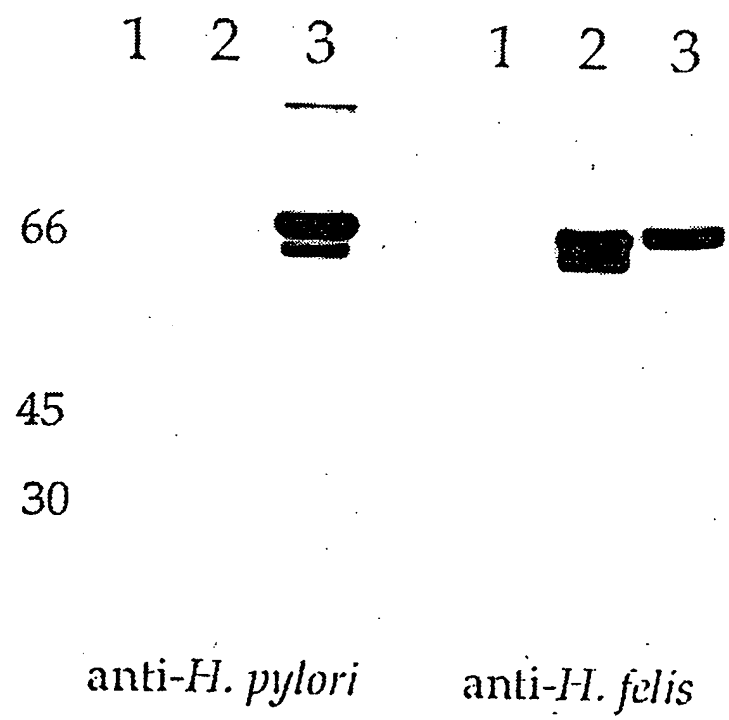


FIGURE 15

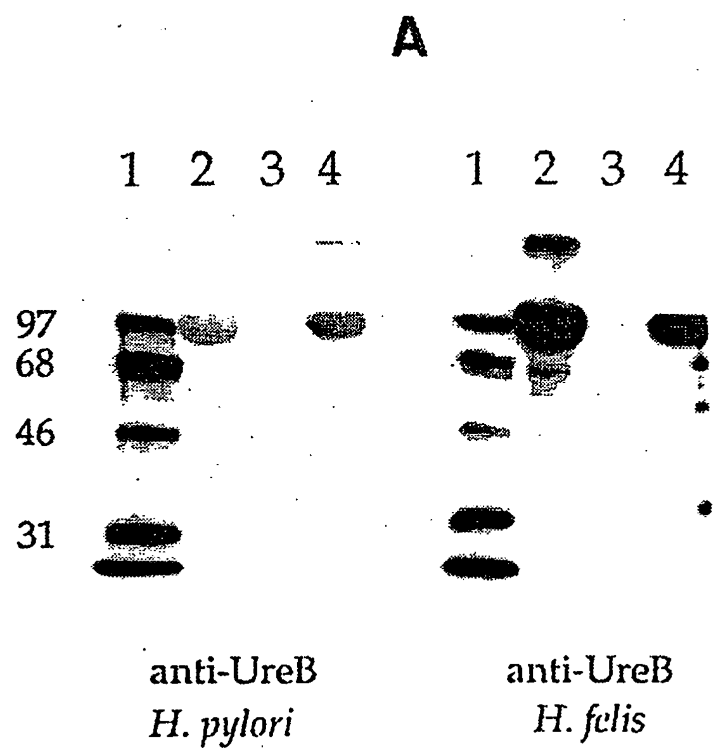


FIGURE 16

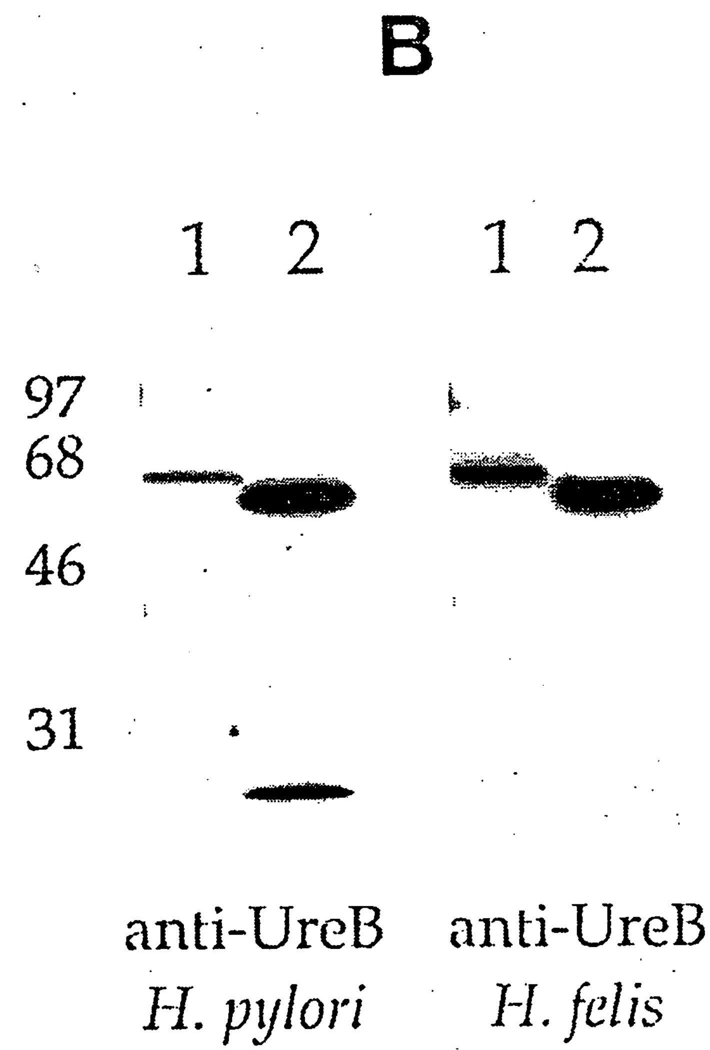


FIGURE 17(i)

HSP A - MBP FUSION

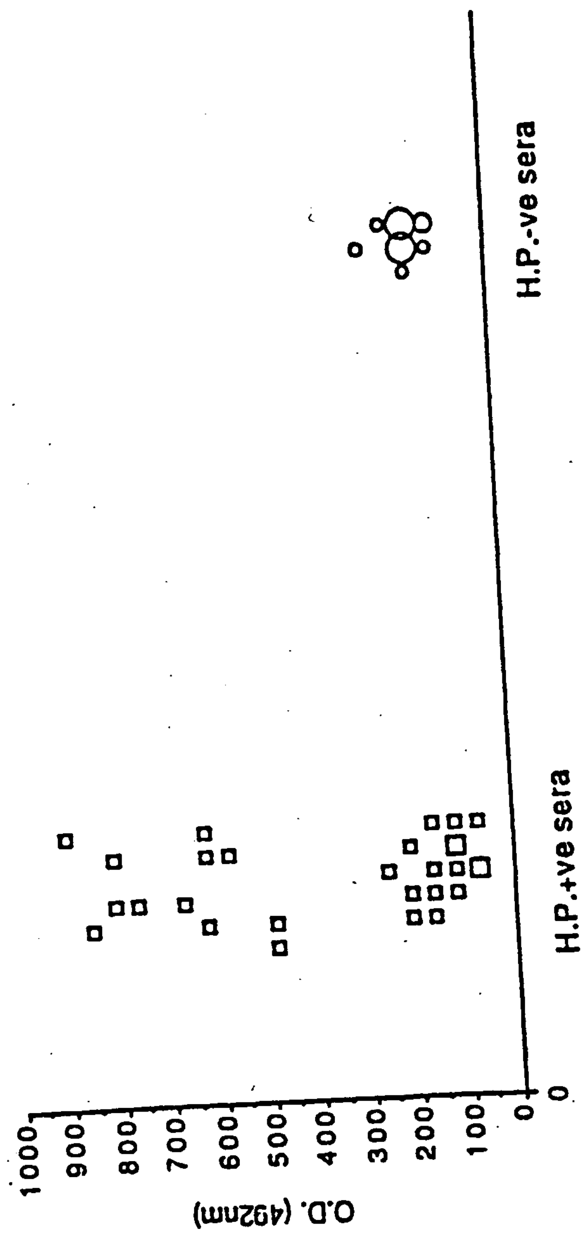


FIGURE 17(ii)

HSP B - MBP FUSION

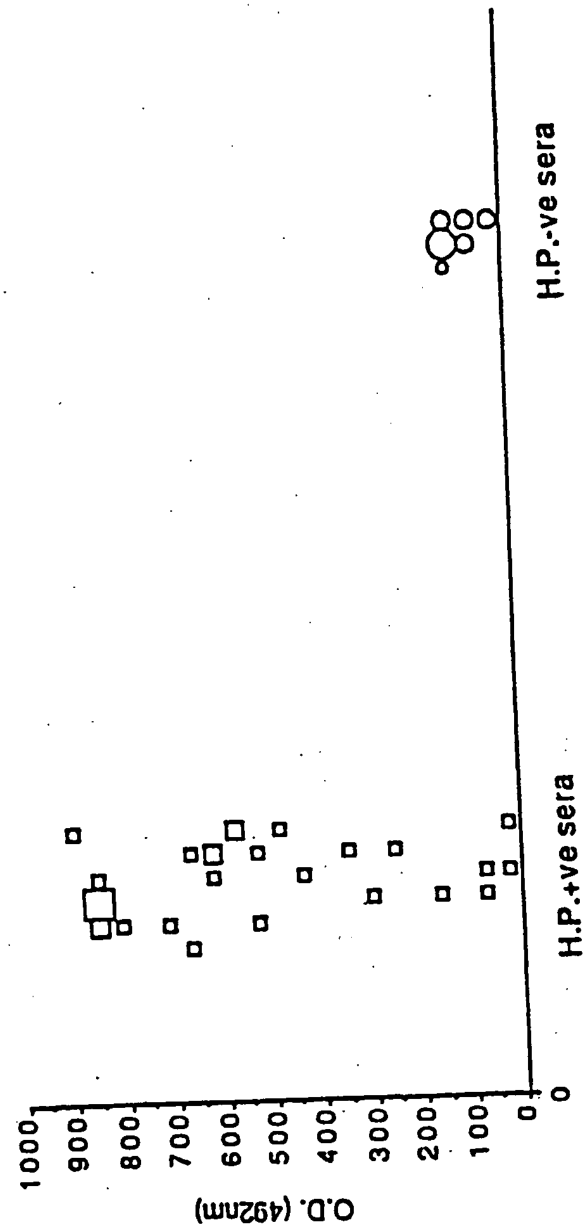
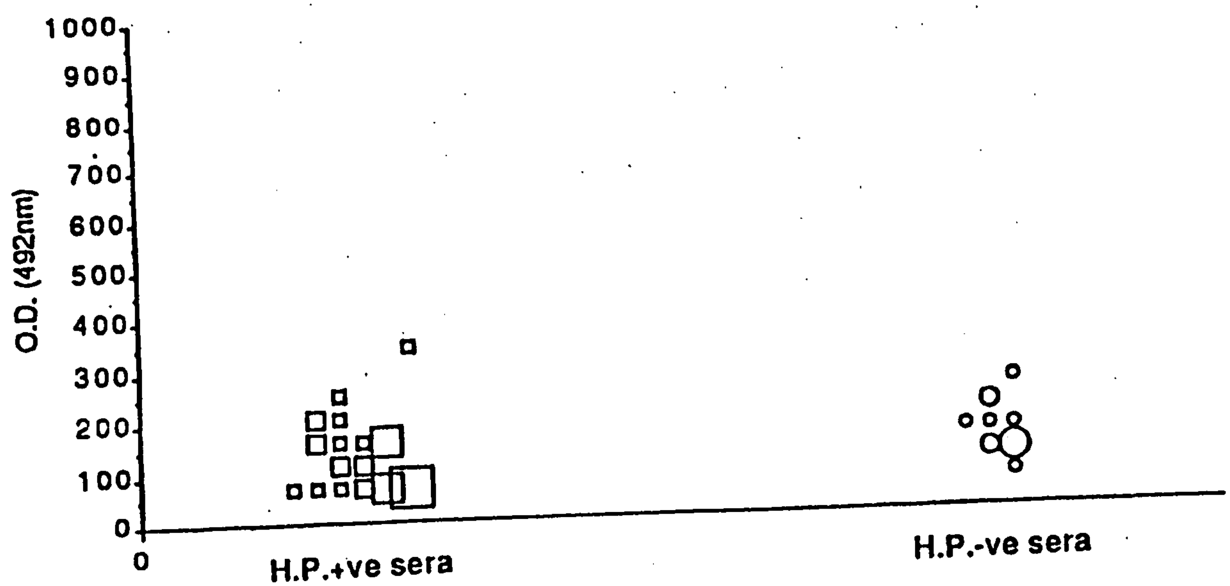


FIGURE 17(iii)

MBP



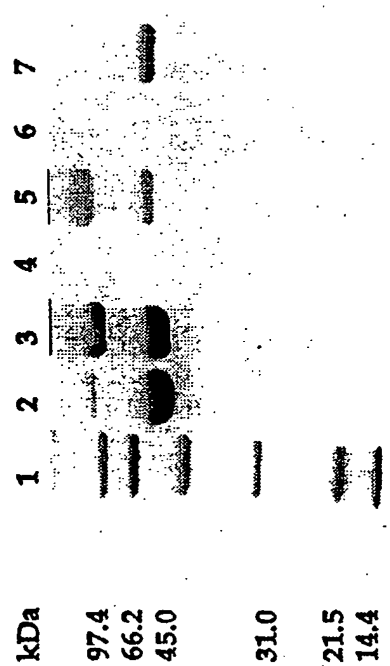


FIGURE 18

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 94/01625

A. CLASSIFICATION OF SUBJECT MATTER
IPC 5 C12N15/31 C12N9/80 C12Q1/68 C12P21/08 A61K39/106
G01N33/577

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 5 C12N C12Q C12P A61K G01N

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	GASTROENTEROLOGY, vol.104, no.4, April 1993, ELSEVIER, NEW YORK, U.S.; page A699 R.L. FERRERO ET AL. 'Molecular evidence demonstrating significant homology between the urease polypeptides of Helicobacter felis and Helicobacter pylori'	10,11,20
Y	Digestive disease week and the 94th annual meeting of the american gastroenterological association, May 15-21, 1993; Boston, Massachusetts, US; * page A699, left column, paragraph 2 * ---	7-9, 12-16, 22, 24-28, 31-35

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

1

Date of the actual completion of the international search	Date of mailing of the international search report
10 October 1994	27. 10. 94

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

International Application No
PCT/EP 94/01625

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO,A,93 07273 (INSTITUT PASTEUR) 15 April 1993 cited in the application * the whole document * ---	7-16,22, 24-28, 31-35
X	INFECTION AND IMMUNITY, vol.60, no.5, May 1992, AM. SOC. MICROBIOL., BALTIMORE, US; pages 1946 - 1951 B.E. DUNN ET AL. 'Identification and purification of a cpn60 heat shock protein homolog from Helicobacter pylori' cited in the application the whole document ---	17
X	INFECTION AND IMMUNITY, vol.60, no.5, May 1992, AM. SOC. MICROBIOL., BALTIMORE, US; pages 2125 - 2127 D.J. EVANS ET AL. 'Urease-associated heat shock protein of Helicobacter pylori' cited in the application the whole document ---	17
Y	WO,A,90 04030 (INSTITUT PASTEUR) 19 April 1990 the whole document ---	10-16, 20,21, 24,34,35
Y	WO,A,91 09049 (RESEARCH EXPLOITATION LIMITED) 27 June 1991 the whole document ---	10-16, 20,21, 24,34,35
Y	J. CLIN. MICROBIOL., vol.30, no.3, March 1992, AM. SOC. MICROBIOL., WASHINGTON, DC,US; pages 739 - 741 P.A. FOXALL ET AL. 'Use of polymerase chain reaction-amplified Helicobacter pylori urease structural genes for differentiation of isolates' the whole document ---	10-16, 20,21, 24,34,35
P,X	WO,A,94 06474 (GALAGEN INC.) 31 March 1994 the whole document ---	27-32, 37,38
P,X	WO,A,93 18150 (BIOCINE-SCLAVO S.P.A.) 16 September 1993	17,20, 22,24,25
P,Y	the whole document --- -/--	34,35

1

Internal Application No
PCT/EP 94/01625

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	MOLEC. MICROBIOL., vol.9, no.2, 14 July 1993, BLACKWELL SCI. PUB., OXFORD, UK; pages 323 - 333 R.L. FERRERO AND A. LABIGNE 'Cloning, expression and sequencing of Helicobacter felis urease genes' see page 324, left column, paragraph 2 - page 326, right column, paragraph 1; figures 1,3 ---	7,8, 10-12,16
P,X	ABSTR. GEN. MEET. AM. SOC. MICROBIOL., vol.93, no.0, 19 May 1993 page 127 S. SUERBAUM AND A. LABIGNE 'Cloning and sequencing of the HSPA and HSPB heat shock protein encoding genes of Helicobacter pylori' 93rd general meeting of the american society for microbiology, Atlanta, Georgia, USA, May 16-20, 1993; abstract no. D-182; see abstract -----	17-26

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 94/01625

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO-A-9307273	15-04-93	FR-A- 2682122	09-04-93
		CA-A- 2120527	15-04-93
		EP-A- 0610322	17-08-94
WO-A-9004030	19-04-90	FR-A- 2637612	13-04-90
		EP-A- 0367644	09-05-90
		JP-T- 3501928	09-05-91
WO-A-9109049	27-06-91	NONE	
WO-A-9406474	31-03-94	AU-B- 4924893	12-04-94
WO-A-9318150	16-09-93	NONE	